PENGUIN MODERN LINGUISTICS READINGS

SYNTACTIC THEORY 1

STRUCTUR[AL(IST)]

EDITOR: FRED W. HOUSEHOLDER

~~8885~~

3671
7.1.87

Penguin Education

Syntactic Theory 1
Structuralist

Edited by Fred W. Householder

Penguin Modern Linguistic Readings

General Editor
David Crystal

Advisory Board
Dwight Bolinger
M. A. K. Halliday
John Lyons
Frank Palmer
James Sledd
C. I. J. M. Stuart

STATE INSTITUTE OF EDUCATION * Govt. of West Bengal

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road

S.C.E.R.T., West Bengal
Date 7-1-87
Acc. No. 3671

1c

Syntactic Theory 1
Structuralist

Selected Readings

Edited by Fred W. Householder

Penguin Books

STATE INSTITUTE OF EDUCATION * Govt. of West Bengal

6880

Penguin Books Ltd, Harmondsworth,
Middlesex, England
Penguin Books Inc, 7110 Ambassador Road,
Baltimore, Md 21207, USA
Penguin Books Australia Ltd,
Ringwood, Victoria, Australia

First published 1972
This selection copyright © Fred W. Householder, 1972
Introduction and notes copyright © Fred W. Householder, 1972
Copyright acknowledgement for items in this volume
will be found on page 331

Made and printed in Great Britain by
Hazell Watson & Viney Ltd,
Aylesbury, Bucks
Set in Monotype Times New Roman

This book is sold subject to the condition that it shall not, by way of trade or otherwise, be lent, re-sold, hired out, or otherwise circulated without the publisher's prior consent in any form of binding or cover other than that in which it is published and without a similar condition including this condition being imposed on the subsequent purchaser

~~8885~~

3671
7 1.87

415
HOU

Contents

Introduction 7

Part One
Syntax and Meaning 21

1 V. Brøndal (1937)
The Problem of Hypotaxis 23

2 D. L. Bolinger (1952)
Linear Modification 31

3 A. G. Hatcher (1956)
Syntax and the Sentence 51

4 V. Waterhouse (1963)
Independent and Dependent Sentences 66

5 W. S. Allen (1964)
Transitivity and Possession 82

Part Two
Categories and Arrangements 91

6 L. Bloomfield (1942)
Outline of Ilocano Syntax 93

7 B. L. Whorf (1945)
Grammatical Categories 103

8 V. Yngve (1961)
The Depth Hypothesis 115

9 C. Watkins (1964)
Preliminaries to the Reconstruction of Indo-European Sentence Structure 124

10 M. Ivić (1964)
Non-omissible Determiners in Slavic Languages 135

11 F. R. Palmer (1964)
'Sequence' and 'Order' 140

Part Three
Relations, Functions and Structures 149

12 Z. S. Harris (1957)
Co-occurrence and Transformation in Linguistic Structure 151

13 A. Martinet (1960)
Elements of a Functional Syntax 186

14 K. Pike (1963)
A Syntactic Paradigm 195

15 P. Healey (1964)
Teleéfoól Quotative Clauses 215

16 D. G. Hays (1964)
Dependency Theory: A Formalism and Some Observations 223

17 R. H. Robins (1968)
Basic Sentence Structures in Sundanese 241

18 M. A. K. Halliday (1969)
Options and Functions in the English Clause 248

19 P. A. Reich (1969)
The Finiteness of Natural Language 258

Part Four
Experiment and Discovery 273

20 P. J. Honey (1956)
Word Classes in Vietnamese 275

21 P. L. Garvin (1962)
A Study of Inductive Method in Syntax 287

22 K. Bergsland (1962)
Morphological Analysis and Syntactic Reconstruction in Eskimo-Aleut 301

23 B. T. Tervoort (1968)
You Me Downtown Movie Fun? 308

Further Reading 318

Acknowledgements 331

Author Index 333

Subject Index 339

Introduction

The choice of Readings for this book is much wider than in the companion volume on transformational syntax. In that volume one may find almost exclusively articles written since 1960 (mainly since 1965) by Chomsky or direct pupils, dealing for the most part with English. For this volume I was able, therefore, to choose anything which did not fit that description, which means almost *everything* prior to *Syntactic Structures* (1957) and a great deal since.

Since 'transformation' has enjoyed three clearly distinct uses in its short career, it is perhaps worth a moment of our time to explain these and note that only the last of the three defines 'transformational' in the current sense:

1. As used by Zellig Harris, first in lectures at Indiana University in the summer of 1953 and later in his LSA presidential address (of which a condensed version is included in this collection), a transformation is a relation (normally symmetrical – actives and passives are mutually related but neither one is prior to the other in any sense) between one set or class of actual sentences and another. The meanings of the two members of such a transformation pair may be the same or different, but are always related.

2. In *Syntactic Structures*, and for four or five years thereafter, a transformation was one of a class of rules which related kernel strings (not actual sentences but only logical precursors of sentences) to actual sentences, and thus was not symmetrical. These rules (some of which were optional, others obligatory) could alter the meaning in certain ways, thus contributing jointly with the lexical entries to the semantic interpretation.

3. In the *Aspects* version (which begins a few years before 1965), a transformation is one of a set of rules which links an underlying or deep structure to a surface syntactical structure, always altering the structure in such a way as to make the semantic interpretation less clear and obvious, but never changing that interpretation (in some versions of the 'standard theory' which have been used since 1968, certain specific details, particularly of the scope of quantifiers, of the interpretation may be due to surface structure, and hence introduced or altered by transformations). In none of these versions is a transformation allowed to make alterations in phonetic substance, except by deletion, repetition or rearrangement (insertion, too, in the earliest version) of whole formatives. Rules which have a more specifically phonetic effect are assigned to a different 'component' of the gram-

mar. Phrase-structure rules (P-rules) develop a labelled bracketing or branching tree-structure; transformational rules alter the trees; lexical insertion rules (wherever placed) and phonological rules affect the substance but not the structure.

The type of transformational article that is excluded from this volume, then, is one which uses the term 'transformation' in either the second or the third way which we have described, either, that is, as a device for altering both meaning and constituent structure in the progress from an underlying form to an actual or surface form, or as a device for altering only the constituent structure (and, indirectly, the phonological order).

What I have tried to do throughout, aided by the advice of many eminent linguists, is to collect the best and most influential (at least one can say this of the earlier items) articles dealing with syntax in the modern era, mostly since 1940. Much excellent earlier material suffers from two fatal defects: it was written about languages which all too few of us even understand nowadays (particularly about Greek and Latin) and for an audience who knew these languages as well as their own and also knew all the grammatical literature about them. These presuppositions are so excessive, that (though I tried very hard) I could devise no way, even with copious annotation, to make any reasonably brief article by, for example, the great Wackernagel or the witty Gildersleeve (or by equally outstanding labourers in the fields of Sanskrit or Romance or Germanic syntax) intelligible to an ordinary modern linguistic audience. And some of the best things they wrote were in footnotes to editions of various literary classics; who ever reads Justin Martyr nowadays even to get the benefit of Gildersleeve's remarks?

What kind of articles are written about syntax depends on a number of factors, but one that should not be forgotten is this: 'What is your definition of the word "syntax"?' So a brief excursus on the changes in fashion related to this point will perhaps be helpful. In large part the choice of definitions is conditioned by whether one is trying to distinguish the domain of syntax from that of morphology or from that of semantics – or from neither, or from both. Though the word occurs in earlier writers, it is Apollonius Dyscolus (very likely influenced by Chrysippus) whose book *Peri Syntaxeos (de constructione)* started the chain of influence into which virtually all later syntacticians fit. What did *he* mean by it? The second paragraph of the first book (translated into Latin by Priscian in the third section of Book 17) contains one of the earliest discussions of the analogy between the two levels of phonology and syntax. Perhaps it is worth quoting in full, with the Greek 'syntaxis' rendered 'putting together'.

> The indivisible matter of the phonemes [apparently 'indivisible' here means 'not further analysable'] accepted this constraint long before [implying perhaps the

logical priority of phonology], which prevents it from making combinations at random, but requires that phonemes be put together according to rule, whence comes the word 'syllable' – almost [etymologically a syllable is a 'taking together']. The syllable itself, the next higher unit, accepted the same restriction, since only when these are put together according to rule do they complete a word. And it is obvious at the next level, the level of words, which are the components of complete sentences, that the grammar here too accepts co-occurrence restriction on putting (words) together. For the senses which are respectively coordinated with individual words (i.e. one to each word, roughly Saussure's *concept*) are the elements of which sentences are made, and as phonemes make syllables in combination, so also the putting together of senses fulfils the syllables by combining them into words. [A puzzling sentence, but Priscian had the same text.] And as the word is made of syllables, so the independent sentence is made from the appropriate co-occurrence of senses.

Notice that it is not just words (as in the earlier use of this analogy by Plato, *Cratylus*, 424E-425A), and certainly not syllables that are put together ('syntaxed') into sentences, but noēta, the senses which belong to the words. This looks like a semantic base hypothesis. But perhaps all he means is that the rules of syntax (sentence construction) do not need any mention of phonological shapes. What are some of the topics Apollonius treated in this book? Here are a few:

The value of the definite article: the correct classification (among parts of speech) of infinitives; various detailed uses of the article; various uses of infinitive and participle; uses of the relative pronoun (known to Apollonius as 'post-positive article'); when verbs take personal pronoun subjects; the accentuation of pronouns and prepositions which govern pronouns; how to resolve the ambiguity of two accusatives with the infinitive; the fact that 'indicatives' are used in questions; the relation of deliberatives to imperatives; which verbs take dative or genitive complements; why prepositions never govern nominative or vocative.

As can be seen from this list, Apollonius does not hesitate to treat matters that one might regard as semantic or morphological or even phonological, rather than syntactic, or to discuss reasons why. The book is a collection of comments on various grammatical topics, roughly grouped by part-of-speech and loosely linked together. The arrangement by part-of-speech still persists in grammars today; what has been added in the meantime (by many grammarians, mostly in the Renaissance and shortly thereafter) are further principles of organization. A nineteenth-century Latin or Greek grammar (there are hardly any brand new twentieth-century ones) is organized as follows (this description is mainly based on H. W. Smyth's *Greek Grammar*): first, phonology and orthography; second, inflection (first nouns, then verbs); third, word-formation (first derivation, then compounding); and fourth, syntax. Syntax is organized into two main parts, the simple sentence and the complex sentence (with a few words on compound sentences in between). In the section on the simple sentence, after a discussion of subject, predicate, concord (agreement), apposition and related topics, the organization by parts of speech begins. In Smyth the order is: Adjective, Adverb, Article, Pronoun, Noun, Preposition, Verb. Most of the space is occupied by Noun and Verb, of course. The section on the noun deals with the uses of the cases (excluding the nominative, which was treated in the introductory portion on subject, predicate and agreement), in the traditional order: Vocative, Genitive, Dative, Accusative (Latin grammars then add the Ablative). The section on the verb deals in order with the uses of the voices, the moods, the tenses, the infinitive and the participle. The treatment of the moods and tenses here is limited to their uses as main verbs. The syntax of the complex sentence is then organized by type of subordinate clause: adverbial purpose clauses, object clauses (other than indirect discourse), causal clauses, result clauses, proviso clauses, conditions, concessive clauses, temporal clauses, comparative clauses, relative clauses, indirect discourse clauses. After this in Smyth come a number of topics which might perhaps have been placed elsewhere, but are regarded as somehow special: questions, exclamations, negation (a complex matter in Greek), sentence particles, tropes and figures.

Notice that the arrangement in the noun is strictly by form (which often necessitates cross-referencing, and results in a wholly inadequate treatment of prepositional constructions – which are relegated to the lexicon), while that for the verb in complex sentences is by meaning and function.

I have described this at some length in order partially to explain the problem which arose when a pattern like this had to be applied to a language like English or French, and later on, when grammarians felt a psychological need for some definition of 'syntax'. Out of the practice of Apollonius or Priscian, or the section on syntax of nouns in grammars like Smyth's came a definition of this sort:

1. Syntax (occasionally called also 'grammar') is the study of the meaning and function of the various inflections (cases of nouns, moods of verbs, etc.) and of the different parts of speech (especially prepositions and subordinating conjunctions).

However, from analogies like that quoted above, from the etymology of 'syntax', and from the section (in grammars like Smyth's) on 'Syntax of the Simple Sentence', a somewhat different definition could be extracted:

2. Syntax is the study of rules for combining words into sentences (or 'into phrases, clauses and sentences').

In a language like Greek or Latin this would be mainly concerned with rules of government or concord involving the cases and moods, but in languages like English – or Chinese, for that matter – it tended to become a study of word-order. So for many linguists research in syntax meant primarily investigations of word-order. This quickly leads research into semantic and stylistic areas, and, because of this, it is often hard to classify a given article.

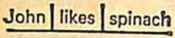
It is this second way, turned around, which led to the name 'part of speech'. Aristotle, at the beginning of the *De Interpretatione*, is clearly still following the lead of Plato (*Sophist*, 262A; *Cratylus*, 425A, etc.), and dividing the *sentence* into two *parts* – subject and predicate, as we should say. But the names given to these parts later became the names of the word-classes noun and verb. The Greek expression which yields our 'parts-of-speech' really meant 'parts of the *sentence*'. So any treatment of parts of speech, as in Dionysius Thrax (the first western grammarian) or as in the recent volume of *Lingua* (*Word Classes*, 1967), is also a treatment of syntax.

The third way of defining syntax is to start from the way Smyth, for example, deals with subordinate clauses.

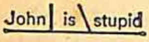
3. Syntax is the study of the devices by which a language expresses the semantic (or logical, or psychological) relations between and among the various parts of a sentence.

Analysis of the sentence into parts (what we might today call 'parsing' or 'IC analysis' – not necessarily 'surface structure' analysis, since many schemes involve permutations and deletions) did not go much beyond the subject-predicate cut in antiquity. The direct object was picked out (e.g. Apollonius, *Peri Syntaxeos*, 3.148), the indirect object (*Peri Syntaxeos*, 3.178, etc.) and certain roles (as we might say) which constitute the so-called categories of Aristotle (*Cat.* 1b: identity, quantity, quality, relation, place, time, posture, state, doing, happening) or the *staseis*, 'statuses' of the rhetoricians (what may be at issue in a lawsuit: the more grammarlike items mentioned – e.g. in Quintilian 3.6 – include usually the first four or six of Aristotle's categories, plus others such as *intention or purpose*, *comparison*, *reason or cause*). The article in this collection by Anna Granville Hatcher represents one modern approach of this kind (so the Greeks often used interrogatives to name the categories or statuses – who? what? how much? how big? what sort? how? how related? where? when? why? etc. – something like the traditional advice to the reporter), and the 'case-grammar' of writers like Fillmore, Gruber and Anderson is another.

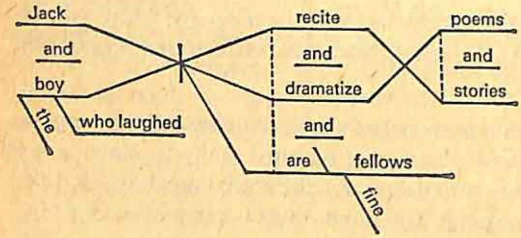
But in medieval and modern times parsing became more methodical and more precisely specified, culminating in the sentence diagramming procedures of A. Reed and B. Kellogg (1877) who seem to have merely modified the scheme of Stephen W. Clark (dating from *Practical Grammar*, 1847) by substituting straight lines below the words for loops around them. There have been other modifications of the procedure since Reed and Kellogg (e.g. the F. V. Irish cited in the *Oxford English Dictionary*), and there were still others[1] between Clark and them, but for at least seventy-five years nearly all Americans (including American linguists) have been exposed to sentence-diagramming at some point between the fifth and twelfth years of school. I have no evidence on the spread of this fashion to Europe, but for those who have never encountered it and for those who have forgotten it, here are some samples.



A single vertical line separates subject from predicate and verb from object. This implies a basic analysis much like Halliday's or Tesnière's.

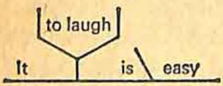


Predicate adjectives or nouns are separated by a 135° (NW) slanting line.



Conjoined items are written on parallel lines; modifiers on slanting lines attached below their heads.

It is easy enough to see that this scheme gives much the same constituency information as a Chomskyan tree or labelled bracketing and the same dependency information as a Tesnière tree. Note that the surface order is not always kept, e.g.,



which quite well anticipates the Rosenbaum extraposition transformation.

However, even schoolbooks which do not use diagrams at all often give complete and detailed instructions for the structural analysis of all kinds of sentences, including explicit rules for ellipsis (deletion-transformations) as well as for transpositions. Although these books do not belong to the linguistic tradition, I would have tried to include a bit in this reader except for the unfortunate fact that, being textbooks, they do not attempt to argue for or justify their point of view, but merely present it as if handed down from Sinai.

---

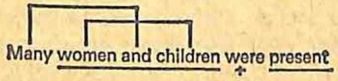
[1] A more nearly one-dimensional system of diagramming is offered by H. R. Greene, *Inductive Language Lessons* (1888), in which modification and conjoining only

Organization on a semantic basis is especially likely to be found when the grammarian is trying to guide the student in his efforts to write, or to write better, in a particular language. So the arrangement in nearly all manuals of Greek and Latin composition is of this kind, as well as in many monolingual composition books. Such a semantic-base approach links up very naturally with a goal which many grammarians over the centuries have accepted willingly though others have spurned it vehemently: explanation. This has at least two aspects: 1. why is such and such an idea expressed in this way? and 2. why is this way better than that in any given instance? Two devices in particular have enjoyed a long and honourable career in explanatory syntax, ellipsis and periphrasis, both of which yield explanations with a strong family resemblance to the 'deep structure' of current theory.

The earliest uses of periphrasis by grammarians (except possibly for some Homeric scholia of indemonstrable age) are in explanations of Greek derivatives. Dionysius Thrax (c. 12), for instance, translates Hyrradios as 'child of Hyrras', and various scholiasts add versions of *Peleides* as 'son of Peleus', *Aeacides* as 'descendant of Aeacus' and others. A little later the scholium of Stephanos (on Dionysius) says (*Grammatici Graeci*, I, 3, p. 225), 'It is criterial for comparatives that they can be analysed into 'more' *(mallon)* plus a positive: sharper = more sharp, slower = more slow'. Philosophers offer some linguistically interesting reductions even earlier. Aristotle tosses off the reduction of finite verb forms to participle plus copula almost absentmindedly (*De Interpretatione* 21b10); possibly the fact that expressions like 'is walking' are not normal in Greek (unlike English), while Greek has nearly as many participles as it has indicative tenses, made this one easier for him. The reference in the *Physics* (185b29)

---

are indicated by linear links above the words, major constituents except the subject are marked by little signs: +, ○, ⊖, ⊕, =. E.g.:



The system used by F. A. March (1869), which he claims to have taken from R. S. Storrs Jr, uses rectangles in place of Clark's loops and avoids slanting arrangements in favour of vertical. He also indicates depth of modification by the number of ticks on the linking line.

leaves the clear impression that this particular analysis had been the accepted thing for some time before Aristotle, perhaps going back to the sophists. It converts verbs into adjectives, so making these two classes distinct only at some superficial level, as in several recent systems.

Another striking resolution of this sort appears in Apollonius Dyscolus, who in three separate passages (*Peri Syntaxeos*, 1.51., Bekker, 31–2, Uhlig, 44–5; *Peri Syntaxeos*, 3.24–5, Bekker, 207, Uhlig, 290–2; *Peri Syntaxeos*, 3.61, Bekker, 231, Uhlig, 326–7) analyses illocutionary performative verbs[2] out of assertions, optatives and imperatives (curiously using a different form in each case: third singular aorist at 1.51 and 3.25 – because he starts out with 'if someone should say . . .' etc.; first singular present also at 3.25, which seems natural to us, and first singular aorist at 3.61, perhaps as self-reportage). The argument is quite modern; these are the verbs that are required as answers to questions of the form 'What did he/you say?' Answer 'He/I ordered you to go (or told you to go)'. 'He/I said/affirmed that he was going' or 'He/I wished it would rain'. When Priscian adapts the last passage (18.47, Hertz, p. 228) he switches to present indicative, first person singular forms, as is natural. We will return to illocutions and performatives when we take up point 4 (see p. 16).

In similar style Dyscolus analyses all first and second-person verb forms as consisting of subject pronoun plus verb, 'us' as 'me and thee' (3.38), and various other phenomena analogously. The short expression in each case is said to be by syllepsis. Closely related is the use of 'transformational' language to account for passives. The rule (*Peri Syntaxeos*, 2.141, Bekker, 174–5, Uhlig, 236, cf. *Peri Syntaxeos*, 3.157, Bekker, 283, Uhlig, 404) goes

---

[2] For anyone who has not followed the recent discussions of illocutionary performatives, a brief explanation may help. Both words are taken from J. L. Austin (1965). Austin speaks of performance of an illocutionary act as 'performance of an act *in* saying something as opposed to performance of an act *of* saying something' and mentions as 'illocutionary forces' things like 'asking or answering a question, giving some information . . ., announcing . . . an intention, . . . making an identification.' Here I am restricting the term to apply to those most universal of illocutionary acts that tend to be marked by special grammatical devices – *asking*, *ordering*, *wishing*, *asserting* and one or two more, and regarding (for instance) *requesting*, *suggesting*, *demanding* etc. as merely variant forms of *ordering*. A performative verb in Austin's use is one which has the peculiar feature that anyone who *says* it (in the first person singular present 'I X' or 'I hereby X') is in fact *doing* it: 'I bid', 'I bet', 'I pronounce you . . .', 'I sentence you . . .', or (with equivalents of the first person) 'This court sentences you' or 'The government sentences you' and the like, where the speaker is speaking for or as 'this court' or 'the government'. The chief simple grammatical illocutionary performatives are 'I ask (you) . . .', 'I bid (you) . . .' or 'I tell (you) to . . .', 'I tell (you) that . . .', 'I wish . . .', 'I hope . . .', 'I wonder . . .' and 'I guess . . .' or 'I suppose . . .'. The discussions alluded to above make the claim that all sentences which do not begin with one of these explicit illocutionary performatives must have at least one 'understood' in the 'underlying form' or 'deep structure'.

'When the verb form is changed to the passive, the accusatives are shifted into the nominative, the original nominative is turned into a genitive with the preposition hypo [by], e.g. *ego se edeira* I skinned you *su edarēs hyp emou* you were skinned by me.' He does not here mention the fact that the personal ending on the verb must be made to agree with the new nominative (subject) since this is a rule of universal application.

Ellipsis has an even longer recorded history in grammar than periphrasis, partly because it can be applied in phonology as well as syntax. In passages near those cited above, Aristotle mentions the ellipsis of *esti* [is] in equational sentences with adjective predicates; this is so common in Greek that it often passes unnoticed. Examples in Apollonius include ellipsis of names of parts of speech when words are cited, as in the phrases *ho* '*men*' (the particle *men*; masculine article to agree with *syndesmos* 'conjunction') and *hē* '*egō*' (the pronoun *egō*; feminine to agree with *antōnymia* 'pronoun', understood). The fact that the distinction between ellipsis and periphrasis is not always clear can be seen in *Peri Syntaxeos*, 2.50 (Bekker, 117, Uhlig, 163): 'In certain words other parts of speech may be contained, as in *Ajax* [a singular; there were two heroes of this name in the *Iliad*, so it often occurs in the dual] the word *one* is present, in *Cronides* [patronymic "son of Cronos"] the genitive singular of the base [Cronos] and the nominative singular of *huios* ["son"]. . . . And no one says that the word *huios* is elliptically deleted with *Cronides* – indeed if it were present the expression would be redundant.' (Priscian, in his version of the passage 17.34–5, Hertz, p. 155, says 'poetic licence' instead of 'redundant'.)

Both types of explanation (ellipsis and periphrasis – including transformations) have continued in frequent use down through medieval and Renaissance times and up to the present day, reaching a climax, by all accounts, in the work of de Sanctis (F. Sanchez de las Brozas), who is said to have derived all uses of the Latin ablative by deletion of appropriate prepositions. For some other often-used devices of grammatical discovery or argument, read the passage in Apollonius Dyscolus where he refutes the view (of Tryphon) that *su* (the second person pronoun) is vocative rather than nominative (*Peri Syntaxeos* 3, 35–42, Bekker, 214–19, Uhlig, 302–308), a view recently revived in modified form by J. P. Thorne ('English imperative sentences' in *JL*, vol. 2, 1966, pp. 69–78).

The fourth way of looking at syntax derives from a passage in Aristotle's *Poetics* 1456b9, where he mentions what he calls *schēmata tēs lexeōs* – perhaps 'sentence-forms' – listing 'command, wish (=prayer), narration, threat, question, reply, etc.' (Question and command occur again in 1457a21–2.) The 'etc.' might perhaps cover the mention (*de Interpretatione* 17a1–10) of affirmations and denials as distinct from wishes (e.g.), though these might be also either 'replies' or 'narrations'. Since wishes and commands are marked by specific verb forms (optative and imperative), the further development of this beginning goes two ways, one toward a theory of moods and tenses of the verb, and the other into an investigation of all the various ways by which these sentence-forms are expressed. And Dyscolus makes exactly this distinction: the moods of a verb are *enkliseis* ('inclinations' or 'modulations'), the 'mood' or type of a sentence is called its *psychike diathesis* ('mental disposition'). The terms are discussed fully by Hahn (1951). The term *enklisis* is taken from Dionysius Thrax who lists indicative, subjective, optative, imperative and infinitive. Some Latin grammarians (e.g. Marius Victorinus) add moods called promissive and concessive, which would seem to be rather either sentence-types (or is promissive somehow equivalent to future tense?) or clause-types. Though we certainly do have performatives 'I promise you that X' and 'I grant that X' in all languages, the fact that Latin had no morphologically distinct optative tended to make morphological moods into *psychikai diatheseis* in the thought of some later grammarians. When we come down to the modern languages, this process goes even farther; the old names are retained (and others added, at times), but they no longer refer primarily to morphological categories; now they are once again sentence-types, illocutions. (This is part of the process that an earlier generation of linguists condemned as the imposition of Latin categories on languages which did not have them.) So here we have still another notion of syntax:

4. What are the illocutionary sentence-types of a given language, and how are they distinguished?

Since modern languages make some of these distinctions with the help of word-order, there may be a partial overlap with definition 2 (or 3). Definitions 3 and 4 both start from meaning, hence we often have difficulty in demarcating the border between syntax and semantics. And since word-order in nearly all languages has stylistic functions, there is often some difficulty in drawing *this* line. Still I have tried in selecting the contents of this volume to exclude items which seemed to me to be solely or primarily concerned with style or with semantics. In spite of that, however, there are examples here of articles whose respective viewpoints may be linked to each of the four views of syntax we have just been discussing.

What have I tried to include? What have I excluded? In the main I have relied on the opinions of many colleagues who nominated various articles for inclusion. I cannot mention them all, but among those who contributed especially long lists are John Lyons, Peter Matthews, M. A. K. Halliday, Robert Austerlitz, Eric Hamp, Frank Palmer, Walter A. Cook, S.J., Howard McKaughan, I. A. Mel'cuk. I have not, however, hesitated to include items mentioned by only one or two people (e.g. the Bloomfield and

Hatcher pieces) if they seemed to me especially deserving, or to exclude items favoured by six or eight if they did not seem to me to measure up. I have tried especially to include articles that are often quoted by other linguists even when (as in the case of Yngve's contribution) many of the mentions are negative in tone.

I had originally hoped to have a more or less continuous sampling from the early twentieth century up to a year or two before the present, but I soon ran into a difficulty which practically eliminates everything before the 1930s. All earlier discussions of syntactical questions (as we remarked earlier) presuppose far too much in the way of detailed knowledge, particularly of Greek, Latin and Sanskrit. These articles could not, as a rule, be made intelligible to a modern linguist without adding a commentary of nearly equal length. In the case of Otto Jespersen there is the added fact that his most accessible works, *Analytic Syntax* and *Philosophy of Grammar*, are now both available in paperback, and no one who reads this collection with pleasure should fail to read those.

A second consideration was the desire to include representatives of each of the major 'schools' or trends in linguistics in this century, as well as from the more notable of the independents. In this desire, too, I was partially foiled, and the following will certainly be missed: Geneva school (H. Weil, for instance), Prague school, Dutch school, mathematical linguistics (especially categorial grammar), Tesnière, recent Soviet work and work in computational linguistics (though Hays, Yngve, Garvin, and Harris have all been connected with such projects). A few of these might have been managed by using extracts from books, but after I reflected on all the problems, it seemed that extending the search to books (especially books currently in print) would not repay the effort. The most important and influential American articles are included in Joos's *Readings* (1967), Bloch (1946), Rulon Wells (1947), Hockett (1954) and Harris (1946). Under other circumstances I would have liked to include all four in this book; but since nearly all our potential readers will already own the Joos *Readings*, it seemed wiser (and gained more space) to exclude these four articles. I *did* get a few nominations of Russian articles, but was unable to make a selection. In picking and choosing, I did not originally have any bias against taking two or more articles by the same linguist. As it turned out, however, there never were two with equal claims to inclusion, so that each author is in fact represented by only one selection. I felt this most keenly in the case of Dwight Bolinger, but there were several names (Hockett, Pike and Halliday, for instance) for whom two or more selections were suggested at least once.

Originally I had thought that I would have difficulty filling a four-hundred-page volume. In the end the difficulty went the other way; there were so many good articles that the task of cutting the list down to size meant many agonizing decisions. The only way to keep X is to cut out Y – or is Y better than Z? Assigning a rank order of merit to thirty or more articles is not easy, and many considerations had to be balanced against each other. But I have no doubt that some of you will wonder why such-and-such an obvious masterpiece is not included, or why this or that clearly second-rate piece of junk is. To you who feel that way I can only wish that you get stuck with a similar job some day. The list for Further Reading includes all items that were nominated as well as those quoted in the selections and a few articles and some books which are intended to make the coverage more complete.

The grouping of the articles is largely on the lines suggested above: first, a half-dozen items with a distinct semantic flavour; next, a dozen with primarily taxonomic bias (the 'part-of-speech' approach mentioned above being one variant of this); third, ten pieces which are concerned in one way or another with relationships, whether paradigmatic or syntagmatic, between sentences or within sentences; and, finally, a half-dozen articles from different viewpoints which treat methodology, showing various ways in which eliciting, statistical analysis, classification and the like can be used to derive adequate evidence for linguistic decisions. Some of the earlier articles (e.g. Hatcher, Harris, Pike) could also have been placed here, and all of these could have been placed elsewhere; Honey's article, in fact, is one which seems to place itself in Part Two. I also considered a separate section for articles with a Comparative-Historical or Typological slant, such as Watkins and Bergsland, but finally decided that this would be less natural and less helpful than the present arrangement.

I think all of us can learn something from the insights of these linguists.

References

APOLLONIUS DYSCOLUS (1910), *De constructione* (*Peri Syntaxeos*), in *Grammatici Graeci*, vol. 2, pts 2–3, G. Uhlig (ed.), Leipzig, Teubner.

AUSTIN, J. L. (1965), *How to Do Things with Words*, Oxford University Press.

BLOCH, B. (1946), 'Studies in colloquial Japanese II: Syntax', *Language*, vol. 22, pp. 200–249.

CHOMSKY, N. (1957), *Syntactic Structures*, Mouton.

CHOMSKY, N. (1965), *Aspects of the Theory of Syntax*, MIT Press.

CLARK, S. W. (1847), *Practical Grammar*, A. S. Barnes & Co.

DIONYSIUS THRAX (1883), in G. Uhlig (ed.), *Grammatici Graeci*, vol. 1, Leipzig, Teubner.

GILDERSLEEVE, B. L. (ed.), (1904), *The Apologies of Justin Martyr*, American Book Co.

GREENE, H. R. (1888), *Inductive Language Lessons*, Lovell & Co.

HAHN, E. A. (1951), 'Apollonius Dyscolus on mood', *TAPA*, vol. 82, pp. 29–45.

HARRIS, Z. (1946), 'From morpheme to utterance', *Language*, vol. 22, pp. 161–83.

HOCKETT, C. F. (1954), 'Two models of grammatical description', *Word*, vol. 10, pp. 210–34.
IRISH, F. V. (1884), *Grammar and Analysis Made Easy*, F. V. Irish.
JESPERSEN, O. (1965), *Philosophy of Grammar*, Norton.
JESPERSEN, O. (1969), *Analytic Syntax*, Holt, Rinehart & Winston.
JOOS, M. (1967), *Readings in Linguistics I*, University of Chicago Press.
MARCH, F. A. (1869), *A Parser and Analyzer for Beginners*, Harper & Row.
PRISCIANUS (1855), *Institutiones Grammaticae*, ed. M. Hertzius, in H. Keil (ed.), *Grammatici Latini*, vols. 2–3, Leipzig, Teubner.
REED, A., and B. KELLOGG (1877), *Work on English Grammar and Composition*, Bobbs-Merrill.
ROSENBAUM, P. (1967), *The Grammar of English Predicate Complement Constructions*, MIT Press.
DE SANCTIS (F. Sanchez de las Brozas) (1587), *Minerva*, Salamanca, J. and A. Renault.
DE SAUSSURE, F. (1949), *Cours de linguistique général*, Paris, Payot (trs. W. Baskin as *Course in General Linguistics*, Philosophical Library, 1959).
SMYTH, H. W. (1956), *Greek Grammar*, Harvard University Press.
STEPHANOS, in *Grammatici Graeci*, vols. 1 and 3, ed. G. Uhlig (1883), Leipzig, Teubner.
TESNIÈRE, L. (1959), *Éléments de syntaxe structurale*, C. Klincksieck.
THORNE, J. P. (1966), 'English imperative sentences', *J. Linguistics*, vol. 2, pp. 69–78.
WELLS, R. (1947), 'Immediate constituents', *Language*, vol. 23, pp. 81–118.
Word Classes, *Lingua* (1967), vol. 17, pp. 1–261, ed. A. J. B. N. Reichling, E. M. Uhlenbeck and W. S. Allen.

Part One
Syntax and Meaning

The Readings in this Part all, in one way or another, continue that ancient practice of searching for a natural or logical explanation of syntactic facts which has led to the development of various forms of Generative Semantics; this is the third notion of syntax in the Introduction. Several of them, in addition, are concerned to explain the underlying relatedness of superficially different phenomena and to establish the existence of syntactic universals of various kinds. Most of them, but in particular Readings 1 and 4, strive to extend the domain of syntax beyond the bounds of the single sentence. Reading 2 is also concerned to explain variations in word-order, a theme which is frequent in Part Two.

415
HOU

S.C.E.R.T., West Bengal
Date 7-1-87
Acc. No 3671

~~8885~~

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road

S.T.I-2

1 V. Brøndal

The Problem of Hypotaxis

V. Brondal, 'Le problème de l'hypotaxe: reflexions sur la théorie des propositions', in *Mélanges linguistiques et philologiques offerts à M. Aleksandar Belic*, 1937, pp. 241–9. Translated from the French by Fred W. Householder.

It is a well-known fact that clauses within a sentence differ considerably in their autonomy or independence in relation to each other. Generally, a distinction is made between parataxis (or coordination) and hypotaxis (or subordination). But there is no unanimous agreement on a clear and unambiguous definition of the distinction.

The problem can be divided as follows:

1. In what sense is a subordinate clause a clause?
2. How can we distinguish subordinate clauses from other kinds (independent, main, coordinate)?
3. Is there more than a single basic type of subordinate clause?

Attempts have been made to define hypotaxis in terms of form, meaning and function. The notion one has of the phenomenon will therefore be, in the first case, morphological; in the second, semantic; in the third, syntactic (or logical).

Morphological notion

According to a widespread grammatical tradition, a subordinate clause is characterized by its form, that is, by the presence of certain words (relative words, mainly pronouns or conjunctions) or of certain grammatical forms (cases, moods). According to a more recent notion, one which is at present generally accepted – a notion which might be called pseudo-morphological – a subordinate clause is defined by its analogy with certain parts of speech – we speak of noun (or substantive) clauses, adjective clauses and adverbial clauses.

Neither of these notions seems to be consistent with all the facts. The pure morphological notion does not take into account the numerous instances in which the subordinate clause lacks any introductory word (e.g. Latin: *rogo venias*, 'I ask [that] you come'; English: *the man you know*). And the pseudo-morphological notion (which should really be extended – why not also admit conjunctional clauses and interjectional clauses?) does not succeed in explaining the purely syntactic difference between subject clauses and object clauses, for example. It is evidently in response to this last objection that a mixed system is often adopted, as by Curme, who distinguishes 'subject, predicate, adjective, object, adverbial clauses'.[1]

In any case one gets the impression that the *morphologists*, who are generally authors of practical descriptive grammars, make no claim to theoretical rigour, and that by using a convenient terminology whose defects they frankly admit, they are trying, in reality, to express essentially independent relations of grammatical form.

We must note here that it is on this highly debatable concept that attempts have been made to establish a historical study of subordinate clauses: from the existence at a certain period of relative words (pronouns, conjunctions) the argument is that one can infer the existence of the corresponding subordinate clauses, and in particular, from the lack of correspondence of such words throughout the Indo-European languages, it is thought possible to infer the non-existence of subordinate clauses in Proto-IE (see especially Hermann, 1894). Such inferences clearly presuppose that a given syntactical situation (in regard to the use of sentences, clauses, sentence constituents) must be linked as a necessary and sufficient condition to a given morphological situation (system of word-classes and morphology). Such a necessary and sufficient relation is utterly inadmissible: this would be a radical misinterpretation of the autonomy of syntactic facts (in the strict sense), it would be to forget the fundamental importance of the clause (and its structure) in every human language.

Semantic notion

Recognizing that in the problem which concerns us the form is influenced by the sense, a recent proposal has been made to base the distinction between parataxis and hypotaxis partly on the difference in semantic independence and partly on the relative psychological importance of the clauses combined. According to Brandenstein, who here takes his inspiration from Marty's linguistic theory, a subordinate clause would be synsemantic (intelligible only in context) while a main clause would be autosemantic (intelligible in isolation) or autonomous from the viewpoint of meaning (1926, 1931). Nehring, though he denies the essentially synsemantic nature of the subordinate clause, nevertheless leaves it a certain element of synsemanticity (1930).

---

[1] The considerable difficulties caused by this doctrine are granted in a very characteristic note: 'These clauses might be reduced to three if we divide them according to the part of speech they represent: 1. substantive clauses which represent a substantive, including *subject*, *predicate*, *object* clauses and such adjective clauses [!] as represent a name in the attributive genitive, or a prepositional phrase; 2. adjective clauses; 3. adverbial clauses.' The author in any case abandons all pretence of theoretical consistency by adding: 'The former classification, however, is for practical reasons usually employed in the following articles, while for the same reasons the latter classification is also at times used' (Curme, 1917, para. 268).

Now we should note here that every clause, even the most isolated (and no clause will ever be absolutely isolated from a context, at least a situational context) is necessarily autosemantic and synsemantic simultaneously – like every word when it is used. The clause is characterized on the one hand by the meaning contributed by its constituent parts and on the other by the part it plays in relation with other clauses. This route, therefore, is not going to lead to a satisfactory definition.

After a closely reasoned discussion of the semantic notion in its two variants, F. Slotty recently reached a conception which is intended to be essentially psychological: what distinguishes a main clause (at bottom the one which is 'determined' or modified) from a subordinate clause ('determining' or modifying) is not any objective, logical difference, but a psychological, subjective one; in the speaker's mind a subordinate clause plays a less important role; it is secondary with relation to the main clause (1936).

What seems disturbing about this psychological notion, or rather the psychological part of this notion – as with the semantic notion in general, in much the same way – is the part which is played by variations which cannot help being idiosyncratic.

Such variations as are marked by sentence intonation, for example, constitute, it is true, unarguable facts, but they are relevant only with certain frames. And it is these frames (which are, in some sense to be specified, logical frames) that are basic for the articulation of thought and, hence, for the functioning of language. Slotty in any case admits this by a lucky inconsistency in establishing his psychological theory on an essentially logical basis (a basis, furthermore, which is highly debatable, as we shall soon see): the psychologically important element would be, according to him, the 'determined' fact (Sachverhalt) expressed by the principal clause, while the secondary element expressed by the subordinate clause would have a 'determining' character.[2]

Logical notion

An old and widespread idea consists in identifying subordination with determination: the 'thing' or 'substance' is considered as more important than the 'accident' or 'quality' and therefore the subordinate clause is defined (with Slotty, for example, as we just noted) as the 'determining' (or 'modifying' or 'qualifying') one.[3]

---

[2] It is worth recalling here that Harald Höffding, the famous psychologist, who is here following the theories of Gabelentz and Wegener, has put forward a theory which is the precise opposite of this: the determining element would be the one which interests the speaker (the *psychological predicate*) while the determined element plays the secondary part of basis or background (*psychological subject*) (Höffding, 1909).

The difficulties with this idea are numerous and very serious:

1. The point of view chosen, that of traditional logic, is exclusively descriptive, arbitrary and inadequate: an object-clause, to take a single example, which is clearly subordinate is in no sense 'determining'.[4]

2. On the other hand, unquestionably subordinate clauses can be found which are so far from being determining as to be, on the contrary, clearly determined – those which perform the function of subject of the main clause.

We must, then, reject the logical notion – at least in this form – as well as the theories which are fundamentally morphological and semantic (or psychological).

Syntactic notion

For a long time a different conception of clausal subordination has been current: one which is equally independent of the form (even the 'inner' form) or presence of words, and of their meaning; that is, subordination as a purely syntactical function. According to Abbé Girard and other excellent eighteenth-century French grammarians (soon followed by German authors) a subordinate clause is the expression of a constituent of a sentence (Girard, 1747). In the same way, since antiquity, logicians (the Stoics, Kant and Ueberweg) admit alongside of independent propositions also some (e.g. hypothetical) propositions which form a part of independent ones. Among contemporary linguists Jespersen has recently defined the subordinate clause in a precisely similar manner: 'a member of a sentence having a form similar to that of a sentence' (1924, 1933, p. 166). This is the notion which I am going to develop here.[5]

To see clearly all the implications of this theory, it is first necessary to distinguish, on the one hand, between utterance and clause, and, on the other, among primary, secondary and tertiary clauses. After defining these terms we will examine the possible combinations of them. Finally, we shall see under what conditions true subordination takes place.

---

[3] This identification of subordination with determination is also fundamental to the theory of syntactic *rank* proposed by Otto Jespersen (1924, pp. 96–7).

[4] In French grammar the term 'complement' (a very vague notion) is often applied to the object. This is an attempt (doomed to fail) to build the entire edifice of syntax on the notion of determination.

[5] Sigwart (*Logic* I, Freiburg, 1889, p. 35): 'Wenn Redeweisen auftreten, in welchen durch Partikeln, Conjunctionen und Relativa verschiedene Sätze verknüpft werden, so geschieht es entweder so, dass vollständige Sätze, die für sich verständlich ein bestimmtes Urtheil aussprechen, in einer Beziehung zu einander gesetzt werden, oder so, dass ein Satz ein integrierender Bestandteil eines anderen Satzes wird.'

Definitions

I By *utterance* we will understand a bounded stretch of speech, autonomous and arbitrarily divided. It may be formed of any number of clauses (and even sentence-fragments).

II The *clause* is here defined as a complete syntactic construction, that is, one which is relatively autonomous from the syntactic viewpoint. It plays no syntactic role in the utterance.

III A *primary clause* is any clause which has no syntactic function within another clause. It may form, by itself, an utterance or a non-syntactic part of one.

IV A *secondary clause* is any clause which functions as a constituent in a primary clause. Several secondary clauses may be simultaneously present in a single primary (or even, in the limiting case, constitute one).

V A *tertiary clause* forms part of any constituent of a sentence. It may enter into a primary or a secondary, or even another tertiary.

Combinations

A Utterances of one (primary) clause.

1. He's sick.
2. Is he sick? Yes.
3. He's staying in bed?! Because he's sick.

Here we may note:

(a) The clause which, by itself, dominates an utterance is necessarily primary; any secondary or tertiary clause which comes to play this role (by a change of expression, for instance) automatically becomes primary.

(b) A series or string of utterances (examples 2–3) is clearly linked internally by psychological relations, since, as Bradley, the great English idealist, says, 'nothing is ever associated without, in the process, being deprived of its independence'. Throughout, there is anticipation or continuation of thought, anaphora, synsemantism more or less close; in replies (2) or rejoinders (3) these relations are particularly close. But in principle and by definition I there are no syntactic relations between utterances.

(c) A primary clause which is isolated and forms an utterance can perfectly well begin with a relative word, pronoun or conjunction (3). Slotty is right to put here analogous examples which were common in the style of pre-war German journalists.

B Utterances composed of heterogeneous clauses (or combinations of primary with secondary or tertiary).

4. He says (that he's sick).
5. (That he's sick) is hardly credible.
6. (If I'm rightly informed), he's sick.
7. (Anyone (who is sick)) must stay in bed.
8. ((No matter) who) can claim to be sick.
9. I said (that I was afraid (that he's sick)).

Here we find clauses, some secondary (examples 4, 5, 6), some tertiary (7, 8, 9), which enter into a primary. Secondary clauses are immediate constituents of the primary (subject: 5; object: 4; circumstance, here condition: 6), tertiaries enter primaries indirectly as a constituent of a constituent (modifier of subject: 7; governing an object within the subject: 8; object within a secondary object: 9).

Two important points appear here:

(a) The secondary – and even more so the tertiary – is by all signs subordinate in all cases. But subordination is realised always and exclusively in relation to the *whole* syntactic superior, not to the non-subordinate (or 'principal') portion of it. So the secondary (examples 4–6) is subordinate in relation to the primary (*including* the subordinate secondary itself). And, analogously, the tertiary is directly subordinate to the entire constituent of which it forms part (the subject in 7 and 8, the object in 9) and indirectly to the entire primary clause.

(b) The ability possessed by non-primaries to play a syntactic role must not be restricted to certain especially obvious cases (subject, object, circumstance). In principle *any* constituent of a sentence may be developed and take the form of a complete clause. The future progress of grammatical analysis which we must hope and wish for will have here an extraordinarily fertile field.

C Utterances composed of homogeneous clauses (combinations of primary with primary or secondary with secondary).

10. (Say), (Peter), (you're sick?)
11. (He says): (I am sick).
12. (He's sick): (he says).
13. (He may be sick), (though he may come anyway).
14. (If he doesn't come), (it's) (because he's sick).
15. (It's Peter) (that's sick).
16. (Is it) (that he's sick)?

In all these cases – obviously very diverse – we are faced by utterances in which the clause-constituents are of the same rank – that is, combinations whose components are not dominated by each other. Some of the clauses are primary (10–12), some secondary (13–16), but anyway they are neither coordinate nor subordinate in relation to each other.

(a) In the case of utterances composed of several primaries one must note that these, since they are primary, are not linked by any properly syntactic relation. It's a case of clauses which by their content or intrinsic value (interjections, imperatives, vocatives, playing the part of a complete member) are particularly apt to function as primaries, or which by their syntactic structure (he says: 11, 12, sentences without explicit objects) require a supplement (what one might call an anaphoric object in these examples).

(b) What characterizes the last set (13–16), composed, according to us, of utterances made up exclusively of secondary clauses, is a balance of elements such as to exclude any mutual subordination (syntactic or logical or even psychological). It is true that in their character of secondary clauses they are all equally and analogously subordinate, but only in relation to the whole utterance. One cannot allow in utterances of this type a main clause with a subordinate as complement. There is only one primary (the only main clause of the utterance) whose constituents are secondaries.

In the four utterances cited in last place we must, it appears, recognize in all cases a subject-clause on the one hand (that is, an expression of a basis or support of a description, or a syntactic *terminus a quo*) and, on the other, a predicate-clause (expression of content or description which is attributed to the subject: 15; which comes here by analysis: 13, 14; where we must hesitantly examine the relation with the subject: 16).

It goes without saying that these clauses (as well as the predicates, in any case) are not of exactly the same sort in the four different types. Analysis will find here an important field for future research.

An argument which is a powerful recommendation of this notion of combined secondaries (a notion which will startle most linguists) to careful consideration is the fact that it was once suggested by the young Leibniz. In his *Generales Inquisitiones* (1686) he hit upon the excellent idea, later rediscovered by the English logician Boole, that there is a perfect analogy between categorical propositions (always primary) and conditional propositions (often secondary), an idea based on the parallelism one must recognize between the analysis of notions (terms) and that of truths (propositions). For example the proposition 'If A is B, then C is D' can be translated as follows:

(A is B) is (C is D)

or perhaps

(A includes B) includes (C includes D) (see Couturat, 1901, p. 254).

Here is the essential principle of true hypotaxis, expounded by the master who was both the best expert on the logical tradition and the inspired forerunner of modern logic. Here, as elsewhere, the theoreticians of syntax would do well to take up the Leibnitzian analysis: it hides unsuspected treasures for linguists as well as logicians.

References

BRANDENSTEIN, W. (1926), 'Kritische musterung der neueren Theorien des Nebensatzes', *Indogermanische Forschungen*, vol. 44, pp. 117–36.

BRANDENSTEIN, W. (1931), 'Erwiderung zu KZ LVII 118 ff,' [see Nehring (1929)], *Zeitschrift für Vergleichende Sprachforschung*, vol. 58, pp. 270–72.

COUTURAT, L. (1901), *La Logique de Leibniz*, Paris.

CURME, G. O. (1917), *A Grammar of the German Language*, Macmillan.

GIRARD, ABBÉ (1747), *Les principes de la langue francaise*, Amsterdam and Paris.

HERMANN, E. (1894), *Gab es im Indogermanischen Nebensatze?*, dissertation, University of Jena.

HÖFFDING, H. (1901), 'La base psychologique des jugements logiques', *Revue Philosophique*, vol. 52, pp. 345–78.

JESPERSEN, O. (1924), *Philosophy of Grammar*, Allen & Unwin.

JESPERSEN, O. (1933), *Analytic Syntax*, Allen & Unwin.

LEIBNIZ, G. W. (1686), *Generales Inquisitiones*, Hanover. Translated as *General Investigations Concerning the Analysis of Concepts and Truth* by W. H. O'Briant, University of Georgia Press, Athens, 1968.

NEHRING, A. (1929), 'Studien zu Theorie des Nebensatzes', *Zeitschrift für Vergleichende Sprachforschung*, vol. 57, pp. 118–58.

SLOTTY, F. (1936), 'Zur Theorie des Nebensatzes', *Travaux du Cercle Linguistique de Prague*, vol. 6, pp. 133–46.

2 D. L. Bolinger

Linear Modification

STATE INSTITUTE OF EDUCATION
Govt. of

Excerpts from D. L. Bolinger, 'Linear modification', *Publications of the Modern Language Association of America*, vol. 67, 1952, pp. 1117–44.

The linear geometry of the sentence imposes certain relationships upon the elements that compose it. These relationships are conveniently expressed as dichotomies. Two dichotomies are, figuratively, spatial: they are CONTIGUOUS versus SEPARATE (A is next to B or is distant from B) and WITHIN versus WITHOUT (A is contained within B or is outside B). Two dichotomies are temporal: FIRST versus LAST and BEFORE versus AFTER. While it does not follow that any of these pairs necessarily has linguistic importance, given their location at the very foundation of sentence structure it would be strange if they had not been seized upon for some use in communication. I think it can be shown that all of them have, probably in all languages, though I am concerned here only with the last two, the temporal dichotomies, and even with these, only in two languages.[1]

Though I assume that the temporal dichotomies probably have some importance in all languages, the particular importance that I should like to attribute to them is hypothetical. For English and Spanish, I believe that the hypothesis has enough footing to be called a theory or even a principle. My reason for not daring to bite off more than English and Spanish at this stage is the extremely subtle nature of the overall manifestation, something that costs a good deal of effort to see just in one's own language, and which it would be hopeless to attempt to trace in any language where one must assume the extra burden of teaching an informant to look beneath the tumble of hills and valleys that make up the surface of his language, for a patch of common ground beneath. Languages are jerrybuilt structures, and the relationships of before and after are boards that lie handy and are likely to be nailed on anywhere. It is an arduous task to determine from the size and shape of the board where it is likely to be nailed and where not. No one could possibly have foreseen, for example, that *He told the squad off*

---

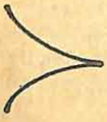
[1] As an example of *within* versus *without*, take the remark of Poutsma (1928, p. 390) that the arrangement *We went to school together* 'could hardly be altered'. We note that *to go to school* is a stereotype with the sense 'to attend school'; but, contrary to Poutsma, we can transfer *together* into the phrase, breaking it up, and destroying the stereotype. *We went together to school* is less likely in the sense 'attended school' and more likely in the sense 'journeyed to school'.

('bawled out') and *He told off the squad* ('assigned') would be in different semantic ranges, while *Let's take the money out* and *Let's take out the money* would be in the same. Side-effects are unpredictable. In many of them sentence order as such has channeled in a certain way the meaning of the stereotype, but usually we cannot determine whether it has or not, since the external features are so gross that our attention is distracted from any possible underlying influence. If we are to find what it is that order contributes of its own, we must seek it in expressions that have the least possible difference in meaning, where stereotyped side-effects are least in evidence. In approaching an informant we must say something like 'I am not asking you for clear and important distinctions, but for hazy and unimportant ones'. So I beseech your willingness to listen for tones that are often elusive, though I expect to offer some that are unmistakable.

The principle

Let us consider what happens when elements – call them words, for convenience – are laid end-to-end to form a phrase. Before the speaker begins, the possibilities of what he will communicate are practically infinite, or, if his utterance is bound within a discourse, they are at least enormously large. When the first word appears, the possibilities are vastly reduced, but that first word has, in communicative value for the hearer, its fullest possible semantic range. The second word follows, narrowing the range, the third comes to narrow it still further, and finally the end is reached at which point the sentence presumably focuses on an event – usually aided by a gesture, a physical context in which only one of several possibilities can be elected, or what-not. This crude picture deliberately oversimplifies the fluid and automatic readjustment that presumably takes place when we ingest an utterance from its beginning to its end. Actually the interaction of situation and spoken symbol is in play from the first; furthermore, in a continuous stream of speech words are not, of course, heard as sharply defined units, either semantically or phonologically. But to the extent that the speaker creates a phrase (that is, to the extent to which the hearer cannot predict a following part of an utterance from what he has already heard) the hearer is taxed to re-shape continuously, and the degree of discreteness or fluidity of the process does not materially affect the argument.

Though circumstances would distort it more or less, we can, therefore, say that the sentence as a complex semanteme has a pointed structure with hyperbolic-flaring sides:



The open end of the horn represents the moment of pre-speaking, followed by the first word and each additional word in course. The narrow end, which focuses on the event, never quite comes to a point, since language is necessarily ambiguous and to sharpen it we must carry it beyond into a non-linguistic setting. But for practical purposes we may represent it as a point.

It follows that words toward the broad end of the horn will be broader in meaning, those toward the narrow end will be narrower, than the same words would be if their positions were reversed. Breadth and narrowness will undergo many transformations depending on the class of words that we are dealing with and on the degree of relevant or irrelevant stereotyping that may have taken place; but with care it is possible to filter out the primary colors from the welter of shades.

An example is the sentence *John eats.* At the moment of pre-speaking, anything is possible; when the word *John* appears this possibility is modified, though John, until we are told otherwise, could be doing anything: the term *eats* then appears and restricts this activity to one kind of act. *Eats* is not a modifier in the terms of Latin grammar, but it just as surely modifies *John* as would the adjective *young* in *Young John.*

The example is unserviceable, however, because it could be argued that *John* also modifies *eats*, as, in a sense, it does, and there is no way of testing the linearity of modification because the order subject-plus-predicate has been stereotyped in this kind of sentence.

It is not, however, stereotyped in Spanish. The Kahanes (1950) found that the type *Juan canta* is used in Mexican Spanish more often than not to signify what John does for a living, whereas *Canta Juan* more often than not means that he is singing now. Actually these particular meanings are secondary, but they derive logically from the fact that in *Canta Juan* it is primarily *Juan* that narrows *canta* – starting off with *canta* the speaker suggests an action that anyone might be doing, hence some such situation as a concert where several persons have been singing and it now comes John's turn. The *canta* is general and the *Juan* is restrictive. On the other hand *Juan canta* suggests possible activities for John[2] – he sings for a living rather than writes or weaves. The one answers the question *¿ Quién canta ?* The other answers *¿ Qué hace Juan ?* Note particularly that the differentiator, the contrasting element, comes last.

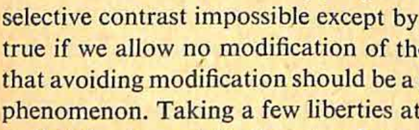
Now for an equally striking example in English. It is obvious that there is a clear-cut difference between *Why did you abruptly back away ?* and *Why did you back away abruptly ?*, though not one so radical as to put them in different semantic ranges. The first asks essentially 'Why did you back away at all?' while the second asks 'Why, having decided to back away, did you do it abruptly?' The first can be contrastive, but only as a whole; we might say *Why did you abruptly back away? Why didn't you courteously accept as I wanted you to?* The second is contrastive on either part, depending on the stress: *Why did you back awáy abruptly when I told you to dart fórward abruptly?* and *Why did you back away abrúptly when I said to do it grádually?* When *abruptly* precedes the verb, it is difficult for contrastive stress to set if off against *back away*. [. . .]

---

[2] Progression is partly crowded out by competition from the *está cantando* construction, which causes *Juan canta* to favor the durative present.

Precisely this same change of meaning cannot be effected by changing the position of the average English adjective, where again we are blocked by a stereotype. *The red house burned down* cannot be worded *The house red burned down*. As might be expected, then, the adjective modifier, unlike the pre-posed adverb modifier, can readily be differentiated by means of contrastive stress: compare the ease of *Why did you choose the réd house?* with the difficulty of *Why did you slówly back away?*

The Spanish descriptive adjective, however, exhibits almost identically the same traits as the English descriptive adverb. In *un hermoso edificio* the speaker has in mind only one building – the phrase is a semantic unit, just like *abruptly back away*; in *un edificio hermoso*, however, the building is contrasted with others that are not beautiful, just as *back away abruptly* contrasts one kind of backing away with another kind. What W. E. Bull calls the 'holotomic' or whole-splitting function of post-posed adjunct adjectives in Spanish is also true of post-posed adverbs in English. When the qualifier precedes, it overshadows the whole of the following noun or verb, as we might expect from their relative positions in the horn of the sentence; when it follows, it splits the noun or verb. Diagramming, we get:



I have said that the stereotyped position of most English adjectives makes selective contrast impossible except by means of contrastive stress. This is true if we allow no modification of the wording, and in the main I think that avoiding modification should be a principle in seeking examples of our phenomenon. Taking a few liberties at this one point, however, will show us that fundamentally the same forces are at work even with the English adjective. Take the example of a man house-hunting in a certain neighborhood. He is asked *Are you looking for something?* and replies *I'm looking for a vacant hóuse*. Or, if he assumes that house-hunting is obvious under the circumstances, he may confirm it selectively by replying *I'm looking for* *a house that's vácant*. Unable to say *a house vacant*, he does the next best thing and expresses the selective contrast by means of an adjective clause. Now why did he not say *I'm looking for a vácant house?* Because contrastive stress on pre-position would rarely be used in such a situation unless *house* had already figured, verbally, in the conversation. Asked *Are you looking for a house?* or *What kind of house are you looking for?* he might readily reply *I'm looking for a vácant house,*[3] but this reply would come, so to speak, at second hand – the sentence would not be of what I term the 'first instance', that is, a sentence that can begin a discourse. [. . .]

I have dwelt upon selective contrast because it is the most striking manifestation of linear modification. It is not the only one, however, as we shall see as we examine other examples of 'free position' in English.

English examples in detail

Simple adverbs

I include here the majority of adverbs except subordinating adverbial conjunctions; I shall not attempt a definition – to do so would require an article in itself.

My example of *abruptly back away* used a descriptive adverb. It was, however, limited to a certain type of question. We shall now see, first, whether other descriptive adverbs in other contexts behave similarly, and second, whether shifting the adverb through a progression of steps creates a semantic gradation.

The selective contrast that we observed in *back away abruptly* is evident in the following example from Poutsma (1928, p. 464), *A man she respected too highly to deliberately tease*, if we end-shift the *deliberately*: *A man she respected too highly to tease deliberately*, the latter arrangement would have given the unwanted implication that she might tease him *unconsciously*, hence the split infinitive. In the original order, *deliberately* modifies all of *tease*. [. . .]

Poutsma points out (p. 439, n.) that certain adverbs of time are 'decidedly the rule' in end-position when the verb is imperative. But he could have included practically all adverbs, since in commands like *Do it carefully*, *Carry them slowly*, *Take it easy*, the adverb is the contrastive point of the command.[4]

If a following adverb so often suggests the possibility of two or more types of the action in question, it follows that a preceding adverb will often suggest only one type of the action. If – and such situations are typical – we already know from the context something about the action in question, the preceding adverb may then be almost expletive, adding nothing to our information. Some examples are: *He was badly shaken (cut, bruised, wronged, surprised)* – such things may be naturally bad. *You are cordially invited, He politely held the door open for me, They courteously stepped aside, He rudely insulted me, I selfishly thought only of myself* – innumerable expressions like these involving the amenities show actions which in our society are normally regarded as right or wrong. By the same token, adverbs of praise and dispraise are highly frequent in pre-position – precisely the same is true of both the adverb and the adjective of praise and dispraise in Spanish.[5] *The thing doesn't automatically stop now* refers to a machine which normally does not stop in any way except automatically; in reference to machinery, expressions like *After the third phase it automatically stops* are common. [. . .]

---

[3] Two different intonations to answer the two different questions.

[4] Commands are typically short. We are not likely to combine two orders in one. Thus in *Do it carefully*, the *do it* part is almost certainly resumptive, *carefully* being the only new datum. If the *do it* is also a new datum, the dual command will probably take some such shape as *Do it, and be careful about it.*

Coming to our second point, whether progression in position creates progression in meaning, we are handicapped more than ever by a variety of side-effects. This is only natural, since given several positions of a word one might expect that some of them would be captured for specific stereotypes or distinct grammatical functions. Take for example the progression *Generally he made himself agreeable, He generally made himself agreeable, He made himself generally agreeable,* and *He made himself agreeable generally.* We can compare the first with the second and the third with the fourth, but not either of the first two with either of the second two (unless there is radical alteration of pause and intonation); this is because the two front positions have been captured by the meaning 'as a rule', and by the function of sentence adverb, leaving the meaning 'in all ways', and the function of modifying a single word, to the two back positions.[6] [. . .]

So we must seek an adverb that is not made captive in this fashion. Again we may use our example with *back away*: *Slowly he backed away, He slowly backed away, He backed slowly away,* and *He backed away slowly.* Selective contrast is unlikely until we come to the third position, and is frequent only in the fourth position. On the other hand, the type of intonation typical of pre-posed adverbs *(Slo-o-owly he backed away)* shades off slightly from first to third position and becomes least likely in fourth position. The third position therefore shares both possibilities – drawling and selective contrast; and there is no mistaking the difference

---

5. *Cortésmente me abrió la puerta* implies '*cortés* to be expected'. *Furiosamente se abalenzó sobre mí* produces the same effect as *furioso león.*

6. There is a further side-effect in the fact that *agreeable generally* admits of two types of selective contrast: 'agreeable in all ways instead of just some ways' and 'agreeable in an overall sense but maybe specifically disagreeable'.

between the first position and the last. We may conclude that gradation of position creates gradation of meaning when there are no interfering factors.

The effect of pre-position is that the adverb colors everything that follows. In *Slowly he backed away* the whole picture, not just some part of it, is in slow motion. I believe that this is significant for what is commonly called the 'sentence adverb'; and I think it suggests that adverbs cannot be sharply divided into sentence-modifiers and word-modifiers, since, depending on its position, and with no side-effects to interfere, an adverb can modify all of an utterance or a unit consisting of almost any major or minor portion.

The adverbs commonly recognized as sentence adverbs would then be those which frequently gravitate toward the front. We have already seen how *generally* not only gravitated toward the front but was made captive there. It does not follow, of course, that adverbs which modify larger units than single words must stand at the beginning: in *Recently he has not been around, to be around* is a unit, and accordingly we may say *He has not been around recently* and still have *recently* modify a phrase. In spite of this, *recently* is more inclusive when it stands at the beginning.

Similar to sentence adverbs are conjunctive adverbs, which have as an additional reason for standing at the beginning the fact that they throw a line to a preceding locution and therefore tend to be as close as possible to it. As with sentence adverbs, an ordinary adverb may be made captive in pre-position in some conjunctive sense. Thus if we reword *I believe it absolutely*, as *Absolutely I believe it*, we get *absolutely* in the sense of 'indeed' and confirming some previous statement. *Yet* is even more striking. In *We yet have nothing to show for it* (I conceive this as possible though infrequent), *We have yet nothing to show for it*, *We have nothing yet to show for it*, and *We have nothing to show for it yet*, there is the gradation that we noted earlier; but in *Yet we have nothing to show for it*, *yet* has been made captive, and to get anything like the former sense we must replace it with *as yet*. [. . .]

Adverbial clauses

These larger blocks evidence linear modification just as clearly as words do. Compare *If you come I'll help you* with *I'll help you if you come*; the first envisages more affirmatively the possibility of the person's coming – has, that is to say, a meaning that goes a little beyond mere condition and may amount almost to an invitation; the second lends itself more readily to the implication 'I'll help you ONLY if you come' – selective contrast. When the *if* clause is placed first, it serves as a frame for all that follows; Professor Hatcher suggests the example *If you drop in this afternoon we'll have tea and play a few hands of bridge*, which proves the point in that it is irreversible:

We'll have tea and play a few hands of bridge if you drop in this afternoon is an improbable sentence. Compare also *When you come I'll help you* with *I'll help you when you come*; the second lends itself to the implication 'but not before' – selective contrast. [. . .]

Participial phrases

When adverbial or quasi-adverbial, these behave very much as adverbial clauses. In *Going that way you'll get lost* the phrase covers the meanings of hypothetical *if*, *when*, and *by* clauses as well as, potentially, the literal implication that the person is actually going or about to go that way. In *You'll get lost going that way* the hypothetical senses more or less fade – the person is probably actually in the process of going that way. Compare also *Cursing his luck he walked out* with *He walked out cursing his luck*; in the first, we tell what else he did while he was cursing his luck; in the second, we tell what else he was doing when he walked out. Samuel Gili Gaya (1943) observes this identical effect in Spanish, comparing it to the effect of transposing the attributive adjective in the type *hermoso edificio* and *edificio hermoso* which we have already studied.[7]

The *-ing* form as a gerund shows a similar peculiarity: in post-position it tends to be limited to the REALITY or IMMEDIACY of the action designated, while in pre-position it is broadened to include the potential as well as the real. Consider *Playing golf is fun*, an utterance which is appropriate when merely imagining the game as well as when actually engaging in it. The alternate form *It's fun playing golf*, however, is less likely to be used unless one is really playing. Later we shall see that much the same is true of adjunct present participles: *a working man* emphasizes the potential, *a man working* the real.

'Prepositional' adverbs

Though it is difficult at times to discern a difference in the types *They left out both pages* and *They left both pages out*, when differences can be discerned they accord with the theory.

First, there seems to be a tendency for the terminal adverbs to be more literal, that is, narrower in their semantic range, while the pre-posed adverbs take on a looser sense which sometimes is so completely transferred that the adverb becomes a captive there in the given sense. *She is going to take her mother out tonight* is made ludicrous in my dialect when worded *She is going to take out her mother tonight* – the latter means 'remove'. [. . .]

---

[7] He says (1943, para. 146) the *gerundio* 'is generally placed after the verb; but it can go before, and in this case the adverbial modification takes on . . . a relief similar to that of the adjective put before its noun'.

Second, post-position is the one that lends itself to selective contrast. Poutsma says (1928, pp. 419–420, n.) it is not 'difficult to account for the arrangement of the following examples, in which the notion expressed by the adverb is contrasted with that indicated by another adverb; thus that of *back* in *Try to force this bolt back* . . ., *off* in *He pulled his dog off*'.[8]

Factitive and causative phrases

These, involving as they do a more or less stereotyped verbal phrase, are much the same as the prepositional adverbs. *I pushed the door open* is appropriate for selective contrast, *I pushed open the door* is not; *This thanklessness will make any further effort on our part useless* is more appropriate for selective contrast than *This thanklessness will make useless any further effort on our part*. Note the resemblance of the latter example to the corresponding adverbs in *Why protest uselessly?* and *Why uselessly protest?* [. . .]

Only

One or two adverbs deserve individual treatment. In *I only want ten*, *I want only ten*, and *I want ten only* there is the same progression that we observed with *slowly*. In the first is a suggestion of 'It is only that I want ten', which is mixed in the second and entirely lacking in the third; yet all three signify fundamentally 'only ten' – the difference is that the pre-posed *only* takes on a broader meaning. The same may be noted in Palmer's example (Poutsma, 1928, p. 447) *I only saw my friend yesterday*, which he uses to show that *only* owes its proper understanding to intonation. [. . .]

So

So may have alternating positions either as a pro-word (pronoun, pro-adjective, pro-adverb) or as an adverbial conjunction.

In '*John is a good dancer*' – '*So he is*' we have *so* resuming something from the previous utterance. Since resumptives set the stage for what is to follow, cast their shadow so to speak over the whole of the coming utterance, we find them often in pre-position. In '*John isn't a good dancer*' – '*He is so!*' we find selective contrast. The greater definiteness of *so* in post-position is also evident in '*They came early*' – '*So you told me*' against '*How do you know?*' – '*They told me so*', and even, though it has to contend with lack of stress, in *It was useless – at least, so they said* against *It was useless – at least, they said sò*. [. . .]

---

[8] The contrary effect may be noted in the manner in which a terminal adverb weakens a contrast on some other element even when that element bears contrastive stress. Thus *Put your hát on* is less effective as regards singling out *hat* than is *Put on your hát*.

Terminal prepositions

Jespersen (1940–49) quotes Fowler as saying that a preposition at the end of a relative clause is admissible 'only when used in the literal or most obvious sense', as in *It was a cab* [but not *high indignation*] *that he rode away in*. It is in accord with our theory that a word placed toward the front should take on a broader meaning; Fowler says in effect that when a preposition includes figurative as well as literal meanings, it must go forward.

With interrogatives we get phrases like *Where from?* versus *From where?* and *What for?* versus *For what?* Take a question-and-answer set like '*He'll get the money*' – '*Where from?*' as against '*He'll get the money*' – '*From where?*' In *Where from?* the *where* is more inclusive: *Where from?* may be skeptical about there being any such place at all, while *From where?* stresses the difficulty of finding the place. The same contrast is evident in *What for?*, where *what* is so inclusive that the phrase as a whole is practically interchangeable with *Why?*, against *For what?*, in which the referent of *what* is pointed to starkly.

With verbs, comparison is more difficult because of the stereotyping of forms like *overlook*, *overrun*. Nevertheless, if comparison be allowed, we note the usual contrast between an expression like *They ran over the field* and one like *They overran the field*: the pre-posed *over* becomes broader.

Simple adjunct adjectives[9]

English has a few adjectives which, through French tradition or some other influence, may stand either before or after the noun – for example, *royal battle* and *battle royal*, *positive proof* and *proof positive*, *payable accounts* and *accounts payable*. Most of them, however, have been captured by some side-effect. Thus *God Almighty* has been spoiled for comparison with *Almighty God* by the fact that it is seldom used except as an oath. *The members present voted* is hard to compare with *The present members voted* because the one means 'not absent' and the other means 'having that status now' (but this may be a derivative side-effect, since the more literal meaning, 'not absent', is the one that follows). [. . .]

Similarly in *If we'll all work and get together the money necessary . . .*, *necessary* gives selective contrast (heightened by the fact that in this order it usually calls for further qualification, like *necessary for our aims*), whereas in *the necessary money* it overshadows and hence binds together the whole. Poutsma likes to point out the resemblance of the first type, *the money necessary*, to undeveloped clauses like *the money that was necessary*. We have already noted the relationship in the case of *a vacant house*.

---

[9] The entire paper, and this section of it particularly, has benefited from the detailed and painstaking criticism of Professor Anna Granville Hatcher of Johns Hopkins University and Indiana University.

There is also marked selective contrast in the example (Poutsma, 1928, p. 489) *It is the man who makes music because he loves it who is the true musician. He is the musician pure*. But this example is revealing in that it shows one device by which English gets around the disability of not being able, ordinarily, to put an adjective after a noun in order to get selective contrast. We note that in the example, *true musician* gives practically the same effect as *musician pure*, and is, in fact, repeated by the latter. The reason is that in *the true musician* (or *the pure musician*) we have re-defined the noun *musician* to exclude all would-be musicians. This gives selective contrast between *musician* and *would-be musician*, which is the same as that between *musician pure* and *musician impure*.[10] [. . .]

Grouped adjectives have more freedom of movement than single adjectives. Thus in Spencer's dictum (Poutsma, 1928, p. 503) *A three-year-old urchin playing with an open razor, cannot be allowed to learn by this discipline of consequences* shows *three-year-old urchin* with its literal significance plus that of 'young urchin in general'; worded *urchin three years old*, the meaning would be pared down to its most restrictive. In *three-year-old urchin* the point is 'urchin'. Similarly in *the three-mile-long passage* the point is 'passage', while in *the passage three miles long* the point is the differential.

Sometimes the accrued meaning resulting from grouped pre-position is emotional or poetic, as may be seen in the rather lyrical *A broad, level and verdant plain* as against the matter-of-fact *A plain broad, level and verdant*, or in the insulting *He is an evil-tempered and malicious man* as against the more matter-of-fact *He is a man evil-tempered and malicious* (The latter paraphrases Poutsma, 1928, p. 498.) Precisely the same is true of Spanish.

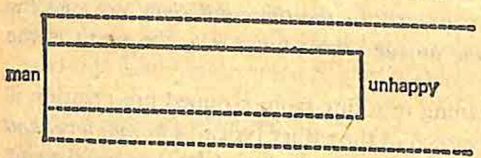
The examples of simple adjunct adjectives that I have cited up to this point seem to have been rounded up from the scraps and left-overs of the English language, and to present no traceable pattern of any consequence. Is there, then, no broad patterning of adjectives that follow their nouns for some effect or purpose that accords with linear modification? If our general principle can aid us in discovering a specific formula, we shall have a test of its usefulness as well as its validity. I believe that there is a pattern, and that examples like *Such countries friendly would be a great asset to us*, *Mary beautiful is something hard to imagine*, *There is nothing more delicious than grapefruit fresh*, *A man unhappy is a social risk*, *Lions are always fearsome beasts but a lion hungry is the terror of the jungle*, *The people sick had to stay in bed*, etc., represent a type that can be defined. To arrive at it, let us enter by the back door, so to speak, along a corridor where post-position cannot be employed.

---

[10] At the risk of laboring the point, I again insert the caution that *He is the púre musician*, with contrastive stress on pre-posed *pure* and with *musician* de-stressed, is iterative.

One of the most important questions about position has to do with the 'standardizing' effect that is produced by placing a qualifier before its noun. Obviously the phrase *underseas craft* names a standard type of craft in our culture, whereas *craft underseas* does not. Linear modification enables us to relate this phenomenon to the more ordinary one just mentioned, *an unhappy man* versus *a man unhappy*, where the notion of 'standard type' is not so evident. The relationship is a real one in English, though my development of it may seem circuitous.

We begin by inferring, from the comparative rarity of the post-position of ordinary adjectives in English, that if a pattern can be found where post-position is regular, that pattern will show a trait of 'restrictiveness' that is more exaggerated than would be found where post-position is common. Now in the type *a man unhappy*, *the people sick*, *the guests absent*, *the tools ready*, *a person busy*, etc., adjectives may regularly follow their nouns. Applying our reasoning to *a man unhappy* we get



in which *unhappy* does not cut off the whole of *man*. Do the parts that it does not cut off represent other men? Let us try a sentence in which there might be a contrast between *a man unhappy* and some other man who is happy: *A man unhappy fell down and broke his leg*. Such a sentence is clearly un-English, whereas *A man unhappy is seldom in control of his emotions* is acceptable English. The parts of *man* that *unhappy* fails to cut off in our diagram therefore do not primarily refer to other men, but to the same man: the contrast is between the man unhappy and the SAME man happy. It is not thing-to-thing, or selective, contrast, but self-contrast. This situation has been reduced to a stereotype with the adjective *proper*: compare *the proper test* (thing-to-thing) and *test proper* (self-contrast); Spanish *mismo* behaves identically. We thus find the conditions of our inference satisfied – not only is there restriction, but the restriction is of a particularly narrow kind; and as we look at other expressions within the same pattern, we find that it is mostly limited to adjectives which can describe varying states of the same individual person or thing – most of the adjectives in the pattern show a striking similarity to participles and adverbs: *unhappy* suggests *dashed* or *disappointed*, *sick* suggests *diseased* or *stricken*, *absent* suggests *away*, etc. So germane is this to English that the language has preserved a small group of adjectives which are stereotyped in post-position and which reveal their kinship to adverbs by having been at one time prepositional phrases: *afloat*, *agog*, *alive*, *apart*, etc.; their reference is to some transitory state. *Overweight* is similar, though occasionally met in pre-position. The same is true of *money galore* – we note the ludicrous suggestion of 'standardization' or 'characterization' if we try to say *galore money*.[11]

I chose the example *a man unhappy* because it is an extreme case, pointing up the unsuitability of post-position to imply much more than temporary state. The example proved that the adjective had to apply to the same man as would its opposite, *happy*, not to other men, and therefore that it could not CHARACTERIZE the man, but could only picture him in a more or less momentary condition. Actually the only essential of the pattern is that the adjective suggests DETACHABILITY FROM THE NOUN – usually the notion of 'that happens to be'; it may at times be used to set thing against thing as well as thing against some other state of the same thing. We can say, for instance, *People unconscious are unable to hear*, where there is self-contrast, or *The only man unconscious was John*, where there is person-to-person contrast; but in both cases the meaning is 'that happen(ed) to be unconscious'. In *Roast beef rare is a dish fit for a king* the detachability is evident. [. . .] We can even use adjectives of a more 'permanent' sort if the combination is something hypothetical, new, or surprising – i.e. if the connection between the adjective and the noun is still loose and uncertain in the speaker's mind. The type *Such countries friendly would be a great asset to us* is more likely than the type *Such countries friendly are a great asset to us* – the hypothetical *would be* stresses the 'non-belongingness' of the adjective and its noun. In a situation where persons of doubtful loyalty are being tested for their allegiance we might find *All persons loyal will be given jobs*, i.e. all persons 'found to be loyal' as a result of the test – the connectedness of adjective and noun is still new.[12] Similarly with *The officers explained their position and all members antagonistic were asked to withdraw*; *Any part defective will be replaced if returned within thirty days*. The implication of newness or unexpectedness is present where adjectives are post-posed after quantifiers like *few*, *only*: *The few men disloyal were discharged*, *The only house vacant is that one*, *No report unsatisfactory will be filed*. The non-connectedness of *Mary beautiful is something hard to imagine* is obvious.[13]

---

[11] Compare *money ready*, a momentary state, and *ready money*, a prearranged class.

[12] If *Citizens* replaces *All persons*, the result is somewhat un-English. See the next section for the reason, *person* being a more inclusive noun.

[13] Non-connectedness is made explicit by certain interpolations: *Mary being beautiful*, *Having such countries friendly*, *All persons found loyal*. It is commonest of all in the quasi-predicatives *I saw John unhappy*, *We found the report unsuitable*, etc. The pattern probably evolved from these types.

If the theory of pre-position for characterization is correct, it should be possible to find nouns whose semantic content admits of characterization by a given adjective, whence that adjective precedes, while other nouns of different semantic content will not, though the circumstances be similar or identical, admit of the same preceding adjective. Take the noun *party*, in the sense of 'one involved in litigation or controversy'. In this legal sense, parties can be CLASSIFIED, in our culture, as responsible and non-responsible, i.e. characterized as such. Therefore we may say *The responsible party will be prosecuted* where we should be far less likely to say *The responsible man will be prosecuted* – the latter, as we have already seen, becomes *the man responsible*, since something as transitory as responsibility for a given act does not normally characterize men as men.

This leads us back to our initial question, why qualifiers that represent a 'standard type' in the culture must precede. Since the qualifier that follows shows a detachable, especially a transitory, state, in order to represent something as fixed, i.e. to CHARACTERIZE (or, better, to STANDARDIZE or NORMALIZE, since characterizations like *all persons loyal*, where the application is in doubt, have been shown to take post-position), we must put the qualifier before the noun. This has led to the creation of odd compounds like *after-the-fact requisition*, *heavier-than-air craft*, *better-than-average person*, *under-the-counter sale*. Note the latter: an under-the-counter sale is one that is marked by certain stigmata; a merchant in totting up the day's receipts might conceivably have such sales in a separate column, which is the same as saying that they are a CLASS apart. *A sale under the counter* does not imply a class, but an incident.

Some terms are needed with which to label the two positions of adjunct adjectives. If we take 'attributive adjective' in the sense of 'one that assigns an attribute', then *attributive* should be limited to adjectives in pre-position. For post-posed adjectives such as have been described, Professor Hatcher suggests *appositional adjective*. We might simply continue to identify them by their position, but it is necessary at least to point out the danger of using the conventional term *attributive* of any but pre-posed adjuncts.

Our pattern provides one clue for the post-position of grouped adjectives. When a speaker combines a series of adjectives he gives the impression of improvising, and improvisation is the opposite of standardization.

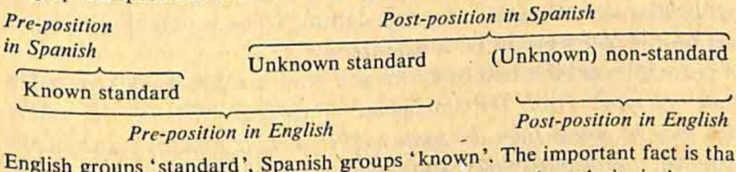
Further evidence for the pattern is provided by adjectives accompanying proper names. In *John unhappy is impossible to live with* we have but one person, John, who has his characteristics that remain unaltered by his being at the moment unhappy; he remains the same John. But to characterize him as an unhappy person we say *Poor, unhappy John!* The *unhappy* overspreads him completely; it is viewed as a trait of his character.[14]

The indefinites

Adjectives with pronouns such as *those*, *the ones*, *anyone*, *no one*, *somebody*, *something*, *somewhere*, etc. do not meet the condition of having more than one possible position; we can only say *anyone wealthy*, not *wealthy anyone* (*a wealthy nobody*, of course, transfers *nobody* out of the indefinites). Nevertheless, they seem to provide either a derivative or a contributing side-effect. Indefinites, by the statement of our problem, ought to precede, and we find that they do; further, there is a bearing on the position of indefinite nouns, where two positions are possible. The chief example is the noun *things*, though other semantically sweeping nouns such as *events*, *times*, *fortunes*, *matters*, etc. occasionally exhibit the same trait: *things incredible*, *events American*, *times medieval*, 'people concerned with *matters Chinese*' (see Hockett, 1951), and so on are phrases which, though somewhat rhetorical, are recognizable English and illustrate our principle of the pre-position of indefinites. Indeed, we might combine this as one and the same manifestation as that set forth in the preceding section. There, it was a highly precise, transitory adjective combined with an ordinary noun which admitted of post-position; here, it is an ordinary adjective combined with a highly inclusive noun. We could say that in either case it is the DIFFERENCE in semantic breadth between the noun and the adjective that determines post-position. The difference in pressure between these two potentials is what counts, and gives us the same feel of 'that happens to be' in *things incredible* as in *man unhappy*.

---

[14] Spanish post-position is not limited to 'detachability', the latter being relegated to a totally different pattern (*ser-estar*), whence 'standard type' is not excluded from post-position. The chief use of pre-position is a special kind of 'norm': that of known or assumed-to-be-known characteristic (see footnote 5 and corresponding text). Just as English post-position is an extreme, so Spanish pre-position is an extreme: that of 'standard' which given a setting is so standard that just mentioning the noun is supposed to suggest the adjective (as in *world everlasting* versus *everlasting world*). Chopping up the spectrum so as to fit the two languages, we get the following:



English groups 'standard', Spanish groups 'known'. The important fact is that while the two languages divide the spectrum at different points, the polarity is the same. Thus all constructions admitting pre-position in Spanish would, on transposing them to the predicative pattern, demand *ser*, though not all constructions in the *ser* pattern when transposed to adjunct pattern would call for pre-position.

Adjunct participles

These, both present and past, have greater freedom of movement than ordinary adjunct adjectives. The example *advancing soldiers* versus *soldiers advancing* has already been found to square with the theory.

The present participles show better than do ordinary adjectives the 'standard' versus 'non-standard' dichotomy that we have just considered. This is to be expected in the nature of participles: they show a process or its result, and a process is impermanent, whence post-position is in the order of things. It is, then, rather pre-position which has special uses, than post-position. The fluid 'characterization' in *advancing soldiers* as against *soldiers advancing* is like that of *sick people* versus *people sick*; but in *traveling salesman* versus *salesman traveling* we have a characterization that has been frozen into a stereotype. It accords with the general theory, for in *traveling salesman* we have 'actually traveling' broadened to 'potentially traveling'.

Such compounds with present participles are numerous, but no different from others with adjectives, adverbs, or nouns plus nouns. It is a characteristic of the type that the qualifying word is transferred from its literal meaning and specialized in some figurative or restricted sense. So the participle partially loses its identity, and functions in the same way in *traveling salesman* as does the adverb *underseas* in *underseas craft*, the adverb *next-door* in *next-door neighbor*, the noun *light* in *light house*, or the gerund *reading* in *reading lamp*. As before, if two positions are possible, the shift will square with the theory – thus though *light house* cannot be compared with *house light*, it is possible despite the shift in semantic range to compare *underseas craft* with *craft underseas* and *next-door neighbor* with *neighbor next door* – in both, the second member of the pair has the more literal meaning.

A non-stereotype is *on the day following* and *on the following day*, between which Poutsma (1928, p. 519) confesses to seeing no difference. Though in a practical situation the two might be interchangeable, they are not identical. *On the day following* implies that time is being measured or fixed according to day periods – *following* contrasts with *preceding* or *same*; *on the following day* offers *day* as a new datum. There is potential selective contrast on the end-word in both examples.

Past participles exhibit best of all the self-contrast that we discovered in *a person unhappy*. *From the standpoint of combat efficiency, a soldier wounded may be worse than the same soldier dead* compares states of the same soldier; *a wounded soldier* would compare soldier *A* with soldier *B*.[15]

---

[15] Although combined with other modifiers there may be a selective contrast, e.g. *The jewels stolen were the ones she prized most of her whole collection*, where *the* gives *stolen* the power to distinguish jewels from jewels, there must always be an implied self-

Other examples of self-contrast: *A pound of coffee hoarded is just that much food likely to go to waste through spoilage*; *A joke misunderstood can cause plenty of trouble.*[16] In these, as in other examples with past participles, post-position keeps the participle more nearly a verb – that is, more literal, narrower, as we expect from our theory.[17] The same is true of grouped participles: *left-over food* versus *food left over*, *a broken-down car* versus *a car broken down*. [. . .]

Similar to past participles and often using past participles are all phrases employing *with* and referring to a position or state – this, of course, is a special case of 'non-characteristic'. *He stood with head erect* (*head high, hands in pockets, arms folded,* etc.) denotes a contrast that is absent in parallel phrases like *erect head* – the latter might be used to describe an animal whose head is characteristically erect, not differentially erect at the moment.[18] The post-positions in this type usually come seriatim.

Special adjectives

Each in pre-position is more inclusive, in post-position more contrastive. *We each received five dollars* might mean 'each' (i.e. togetherness) loosely in respect of sharing an experience, which could be at any time or place. *We received five dollars each* almost necessarily refers to but a single transaction – the 'each', or togetherness, is sharpened to a single occasion. Note that *What a coincidence! We each received five dollars on the same day!* is acceptable, but *We received five dollars each on the same day* is less so because redundant.

More, *extra* and *additional* function alike. In *He gave me a dollar extra* the point is *extra* – we are talking about an enlarged AMOUNT. In *He gave me an extra dollar* the point is precisely *dollars*.

---

contrast as well. The same 'jewel from jewel' distinction is present in *The stolen jewels were the ones she prized most of her whole collection*, but the example changes because of the missing 'previous state versus present state' contrast.

[16] Professor Hatcher notes that these have an element in common: they are gnomic predications with generic subject.

[17] Pre-position assimilates participles to ordinary adjectives; post-position assimilates ordinary adjectives to participles. Note the 'adjective' sense of *C and H is the best refined sugar* against *C and H is the best sugar refined (fallen woman—woman fallen; a given number—a number given; a studied remark—a remark studied)*. In *He is an easily-satisfied person* the participle is not only adjectivized but broadened to suggest 'easy-going', 'friendly', 'comfortable', etc. *He is a person easily satisfied* is verbal and restricted.

[18] The potential difference is not always actualized; thus *with arms folded* and *with folded arms* do not show the striking difference of *head erect* and *erect head*. Its absence is an accident. On the one hand, there is no creature characterized by folded arms as there are animals characterized by erect heads, so that a radical difference in meaning is impossible. On the other hand, there IS a standard pose 'folded arms', which therefore admits the 'standard' position of the adjective.

Intensive pronouns

I myself seldom do it that way is diffuse, including practically everything that may be subsumed under 'as far as I am concerned in the act'. In *I seldom do it that way myself*, however, we focus a selective contrast on *myself*, conceding the right of others to act differently. The type *It was signed by John himself* is contrastive even though no other position is possible.

Nouns plus nouns

During the Second World War combinations of the type *Operation Pacific*, *Operation Claire*, etc. gained currency. The greater particularization of *Operation Pacific* over *Pacific Operation* is obvious. The invention gives us an opportunity of observing a 'title' in the making; for if it continues, *operation* in this construction may take the same status as *king* in *King James* or *number* in *Number Ten*. A recent book on civil liberties is advertised as *Operation Gag*.

Predicate complements

Complements of *be* are striking. In *Here is the money that you requested*, *here* is broadened to include proffering as well as location; in *The money that you requested is here* one detects only literal location. In *Gone are the times when . . .*, *gone* overspreads mere *times* to suggest all that is bygone, while *The times when . . . are gone* is strictly limited to particular gone-times. [. . .]

In '*Why do you want it in money rather than goods?*' – '*Money I can hide, goods I can't*' we have a resumptive, which, in common with other resumptives, overshadows what follows when placed in this position. It is as if we said 'As for money, I can hide it.' [. . .]

Interplay of forces

At various points we have seen positions interlocking with stress and intonation to produce a certain effect. Lest we fall into the temptation of expecting to find one set of variables conditioned by the other, let us compare certain aspects of stress with what we have discovered about linearity. [. . .]

Poutsma says: 'When an adverbial adjunct is placed between the verb and its object, this is mostly done owing to its being subservient to the other elements of the sentence, especially the object itself. . . . Compare *He heard again the noise on the stairs*, and *Again he heard the noise on the stairs*; and observe the relatively weak stress of *again* in the first example' (pp. 416–17, n.). But *again* is 'weak' in this position only if made weak; I find it easier to weaken a terminal *again* than one standing anywhere toward the front.

If stress has any automatic tie-in here it is more likely with the word *again* than with any feature of position; replace *again* with *sometimes* and the word can be weak in all positions.

We have seen how linearity establishes selective contrast, and how stress operates now with it, now across it, for its own special types of selective contrast. Thus in *Like strains of long-forgotten músic* we may re-word to *Like strains of músic long-forgotten* in order to get selective contrast on the adjective and at the same time keep *music* a new datum; whereas if we retain the original order and stress the adjective we get a selective contrast that is iterative: *Like strains of long-forgótten music.*[19] Consider now an instance of selective contrast in which stress assumes the entire burden, the modifier going not at the end nor anywhere near the end of the sentence, but at the beginning – and going there because it has to by the principle of linearity:

In answer to the question *Do you see her Tuesdays?* one may reply *No, Tuesdays I stay home*[20] or *No, I see her Mondays.* The first order comes by reason of *Tuesdays* being resumptive, the second by reason of *Mondays* being contrastive and going where such contrastive elements go, at the end. We would not ordinarily find a combination like '*Do you see her Tuesdays?*' – '*No, Mondays I see her*'.[21] But observe now that a combination like '*Do you see her Tuesdays?*' – '*No, Tuesdays I stay at home; Mondays I see her*' is as normal as the same discourse with the last sentence changed to *I see her Mondays.* There is a selective contrast with the variable at the beginning of the sentence.

This is because in the order *Tuesdays I stay at home, Mondays I see her* we have an ENUMERATIVE CONTRAST, and it develops that any noun or 'substantive adverb' (an adverb that names a place or a time) may be so employed. So combinations like the following are common: '*Tell me about your plans*' – '*Well, tomorrow I intend to work, and day after I'll go fishing*'; '*What's your work schedule like?*' – '*At home I put in three hours, at the office six*'; '*Tell me about your purchases*' – '*This couch we bought at auction, those chairs at a store.*' In each instance the pre-posed element is a logical subject which does not primarily limit what follows, but is limited by what follows, and hence stands first; but which stands in contrast with some other logical subject (not necessarily resumptive) and so receives contrastive stress. The key-word, the overshadowing word, in each such instance stands first, just as in the succession *First we write, second we read, then we play*.

---

[19] The same effect may be observed in Spanish with *Juan canta*, putting contrastive stress on *Juan* so that *canta* is resumptive. In *Canta Juan* both elements are new data, unless *canta* is given contrastive stress, in which case *Juan* is resumptive.

[20] Or, as with sentence adverbs in general, put *Tuesdays* at the end following a pause.

[21] No more than we would find '*How did you do it?*'—'*Carefully I did it*'.

I have not attempted to give more than a sample of the effects of mingling stress and linearity. It would be just as difficult to do more than this as to attempt to describe all the possible shades that can be got by mixing two primary colors – even if we had identified the colors, which, with stress, we have not.[22]

---

[22] I suspect, though the two do not automatically condition each other, that linear modification is one reason for a primary feature of the English sentence, namely, the tendency toward loud stress at the end. If the 'point' of the sentence is there, so will the stress be, more often than not. The same is true of Spanish.

References

BULL, W. E. (1950), 'Spanish adjective position', *Hispania*, vol. 33, pp. 297–303.

GILI (Y) GAYA, S. (1943), *Curso Superior de sintaxis Española*, Mexico City, Minerva.

HOCKETT, C. F. (1951), 'Review of *Nationalism and Language Reform in China*' by J. de Francis, *Language*, vol. 27, pp. 439–45.

JESPERSEN, O. (1940–49), *Modern English Grammar on Historical Principles*, Munksgaard.

KAHANE, H. and R. KAHANE (1950), 'The position of the actor expression in colloquial Mexican Spanish', *Language*, vol. 26, pp. 236–63.

PALMER, H. E., and F. G. BLANDFORD (1939), *A Grammar of Spoken English*, Heffer.

POUTSMA, H. (1928), *A Grammar of Late Modern English*, P. Noordhoff.

3 A. G. Hatcher

Syntax and the Sentence

A. G. Hatcher, 'Syntax and the sentence', *Word*, vol. 12, 1956, pp. 234–50.

STATE INSTITUTE OF EDUCATION
Govt. of

Of the criticism which, in recent years, has been levelled against the mentalist scholar in linguistics, it could be said, as far as descriptive syntax is concerned, that his worst flaws (i.e. those inherent in his method) have mainly passed unnoticed, while he has been severely blamed for his virtues. In my opinion, he has, at his best, two great virtues, and, even at his best, two great flaws. The virtues are, first, his stubborn refusal to abandon the most important and the most difficult aspect of syntax: the search for more specific relationships between form and meaning in language; and, second, his ability to make 'subjective' judgements, to react personally to language: maintaining, that is, a connection between the conscious, analytical part of himself that represents the investigator, and that part of his nature which is sublimely unconscious of its own complicated logical processes and which represents the creative source of language. It is only when this source is tapped that any discovery for syntax can be made.

Nevertheless, the results so far produced by our 'tapping the source' have been woefully inadequate from a scientific point of view; the often brilliant intuitions (e.g. of a Jespersen) that throw a new radiance over the construction being studied are never definitive, and at times the number and variety of the exceptions are such as to make one doubt the validity of the 'tendency' that has been isolated. This great difficulty that we experience when we seek to explain linguistic reality in terms of our intuitions can be largely explained by the twofold flaw inherent in our mode of working and of thinking; the fragmentariness of our approach and the unsystematic nature of our categories.

If I praise the mentalist for his ability to appeal to subjective criteria, but damn him for the unsystematic nature of his criteria, this means that I do not assume (as some scholars appear to do) that 'subjective' and 'unsystematic' are synonymous. Our traditional syntactical categories, for example, largely the product of ancient and medieval grammatical thinking, are relatively objective, but it would be difficult to imagine anything more unsystematic than this unwieldy conglomeration of quite different sets of distinctions: one cannot go from the preterite tense to the indirect object, to the superlative degree, to the inverted subject: each time, one must back out of one blind alley and jump over to the entrance of the next one.[1] On the other hand, it is quite possible to imagine an absolutely systematic classification of (what appear as) intangible nuances: what else is the color spectrum? Nor is the physicist in the least embarrassed that 'red' is always melting into 'purple' and into 'orange'. He knows that what we call 'red' must melt, and why; he knows what it will always melt into on the 'right' and on the 'left', and what these colors, in their turn, will melt into. The basic flaw of our semantic distinctions has been not that they are subjective (my 'red' being your 'purple') but that, when we speak of a certain nuance characteristic of a given construction and must confess that it tends to give way to other nuances – we do not know why this is so, and we do not know the relative positions of the various semantic nuances on a scale: even when the color of our meaning is relatively pure we cannot really see it, because we cannot see, at the same time, all the other chromatic possibilities that are given. (If we could, we would then be able to dispense with our approximative color-words, as the physicist has been able to do.) It is a truism that we can understand a thing only when we can assign it its exact place among all other comparable things, but in our semantic analyses of constructions, we have been attempting the impossible task of trying to isolate this meaning or that meaning or another meaning, without possessing any system of classification of meanings.

And we have been fragmentary in our approach, we have been dealing with parts instead of wholes, precisely because we have sought for meaning only in our constructions. Constructions do not exist, in speech, except as parts of predications; these are the wholes with which syntax should mainly have concerned itself, and our concern for meaning should have brought us to consider, first of all, sentence-meaning. 'What is the meaning of the passive?' is, at the present stage of our knowledge, a nonsense-question; we must ask rather 'what is the meaning of *the stockings were hung by the chimney with care*?' This meaning, whatever it is (and it is surely different from *the stockings were left behind*),[2] must be the same whether one is studying the passive, or the simple versus compound tense *(were hung)*, or the position of adverbial elements *(by the chimney, with care)*, and it must obviously be conceived in terms that transcend voice and aspect and word order. This sentence can be studied, and should be studied, without constructional prejudice, as a unit in itself.

---

[1] It would appear that certain structuralists, perhaps misled by the systematic patterns to be discerned in phonemics, have failed to recognize the totally unsystematic nature of syntactical or morphological constructions (as a whole), pointing out, here and there, certain dim correspondences between the disparate sets of distinctions or their parts. But 'here and there' has no place within a system.

[2] In anticipation of my system of labels, I shall call the meaning of the first sentence, Z; that of the second, O.

It is true that 'consideration of the meaning of the sentence' is a tenet we mentalists cherish, but our attempts to apply this have been sporadic and, usually, vague. It is only seldom that we actually refer to the point or purpose or theme of the sentence as a unit and, even then, we cannot be very precise. To really understand the meaning of a sentence we must understand its relationship to all other possible sentence-meanings, but on the plane of the sentence, as also on that of the single constructions, there exists no set of criteria, applicable to all utterances of a language, to which we may appeal. In the absence of such a logical system it is not surprising that we fail to attain or even to approximate that desideratum of all scientifically-minded scholars: the ability to make definitive, useful statements; to predict in the individual instance.

If, as was suggested above, these two great flaws have passed unobserved by our enemies, the reason should be obvious. How could the mechanists who, in their study of constructions, appeal to the totally unsystematic criteria of Form, chide us for our lack of a system? And our partial disregard of the sentence as a unit becomes with them a total disregard. As for our failure to offer definitive statements, the statisticians, under the label 'Relative Frequencies', exhibit their wholesale failures as scholarly results.[3]

We are all fragmentarians and all hopelessly unsystematic, but with one important difference: the mechanists, who are completely the slaves of Form, can never find a way out of their criss-cross of blind alleys, being doomed forever to the meaningless task of redistributing and recombining the odd-shaped parts of a non-systematic totality; we who seek for meaning, however, can considerably improve upon our methods for attaining it – if we establish ourselves upon the plane of the sentence and, there, seek to construct a set of semantic criteria for the analysis of sentence-meaning.

I believe that the construction of such a system is possible, and in the remaining part of this paper I shall offer a schema (abbreviated, and pre-

[3] In my criticism, implicit or explicit, of the exponents of 'scientific linguistics' who have invaded the field of syntax, I am most directly concerned with those (represented mainly by the statisticians of the Keniston School) who, in their simple, mechanical way, are nevertheless interested in studying constructional usage on the basis of concrete linguistic data in a given language. As for the others, who find ways to occupy themselves while ignoring this problem, or ignoring concrete data, or even ignoring 'natural language', I would say only that their busy, varied activity is largely a by-product of the failure to explain constructional usage, either in terms of meaning or in terms of form. If it were possible to believe that in any language one could arrive at a definitive (or near-definitive) description of a single great construction by reference to actual usage – no serious scholar in syntax (and in such a situation it might be possible to identify the genuinely serious scholar) would be tempted to such extravagant futilities.

sented in rather vague terms) which may serve, at least, as a beginning. In a monograph [*Word*, vol. 12, 1956] devoted to Spanish word order, I have attempted to show how this schema may be applied to given problems in a given language; how it may serve in the analysis of constructional usage. For, when I urged that we begin by forgetting about the meaning of constructions, thinking rather in terms of sentence-meaning, I did not intend to imply that the search for constructional meaning should be abandoned: it is our ultimate goal. But this meaning can be found only in terms of the sentence: only with predications is it possible to escape from our blind alleys, only with predications does there exist a continuum of meaning, in the expression of which the various, disparate constructions play their combinatory role.[4]

I shall begin by stating that the system of classification with which I shall be primarily concerned, based upon one main criterion, is only one of several possible semantic systems: I can think of at least three quite different criteria (and there may be more) to which one might appeal in the semantic classification of sentences: 1. reason for utterance; 2. theme; 3. point of view.

To me the first presents the greatest difficulties. There is, to begin with, the problem of self-expression versus information; granted that we could agree on those utterances representing the latter, there would still remain the very difficult task of classification according to the types of information that the speaker is trying to offer us. In a purely verbal context the problem might be stated: 'what is the relationship, if any, between this predication and the one that has preceded?' – leaving aside, for the present, those cases in which no relationship can be discerned: e.g. brusque changes of subject. But even when continuity is observed, I find myself able to see clearly and sharply only the distinction between 'temporal sequence' (whether narrative: *and then he left*, or instructional: e.g. cookbook recipes), and those predications which do not take us forward in time. Among the latter, however, I have been able to note, in quite confused fashion: descriptions, explanations, illustrations, corrections, objections, corroborations, deductions, conclusions, personal appreciations – and what I can call only

---

[4] The statement 'only in terms of the sentence can we find the meaning of constructions' is over-bold but, at the present time, I am not sure precisely what modifications should be introduced. Perhaps this much may be said by way of tentative modification: assuming that constructions can be divided into those that involve nexus and those that involve junction, it is in regard to the former that my statement (constructional meaning depends upon sentence-meaning) has the greater chance of being true, and useful; with the latter I should like to believe in the ultimate relevancy of sentence-meaning, and I shall give an example later (footnote 20) to show that this may not be an impossibility.

‘information proper’ (or ‘further information’). Everything remains to be done here; I believe that a classification of types of information, in terms of intention, would be, in its partial way, of inestimable value for syntax (and, if we once understood ‘Information’, we might come to understand ‘Self-Expression’). But here I have nothing to offer.

As for Theme [*Word*, vol. 12, 1956] the first chapter of my monograph is devoted to this problem – or rather, to the description of a single, very basic theme, found everywhere in Spanish, everywhere in the languages with which I am familiar: that of the Existence of the Subject *(there is a mosquito, then John appeared)*. Even in that study, however, I have not defined a theme, much less outlined a system comprehending all possible themes; the themes of language can be discerned only gradually – as we come to glimpse them through the interstices of the network of the third system, which I shall now describe: that of Point of View.

In attempting to establish the point of view of a predication I proceed on the assumption that any statement may be considered as the answer to an underlying question, which means that I must attempt to construct a set of questions which might account, in their way, for all predications; the basic criterion of such questions I choose in the distinction between the relatively known and relatively unknown elements of a predication: e.g. with *what happened?* nothing is known (except the possibility of a happening); with *what did he see?* everything is known except the object.[5]

In applying the criterion of the known versus the unknown, one may stress either the first or the second; I shall base my classification on the unknown element (this is necessary if we resort to questions). Assuming the main elements of the sentence to be Subject, Verb, Object,[6] this gives us immediately three (sets of) questions: *Who (what) is the subject?*, *What (who) is the object?*, *What is the activity?* It is the last which is the most complex, for here, there are four possibilities of givenness: Zero is given (Z), Subject is given (S), Object is given (O), Subject and Object are given (S, O):

---

[5] The distinction between the higher or lower informative value of the elements of a sentence is a commonplace to all those interested in Aristotelian logic, including a number of scholars in syntax, nor is the appeal to question-and-answer, in establishing the point of a predication, quite unknown to grammarians. Thus Le Bidois (1952, p. 350) explains the inverted subject with *rester* by citing Gougenheim's analysis of Bossuet's famous sentence: *Restait cette redoutable infanterie de l'armée française*: ‘. . . cette phrase ne répond pas à la question: «Que faisait cette redoutable infanterie? – Elle restait», mais bien à celle-ci: «Que restait-il? – Cette redoutable infanterie»’ (1938, p. 112); this interpretation is also quoted by Harmer (1954, p. 76).

But of the many hundred examples of inverted subject cited by Le Bidois, it is only with the one verb *rester* that he appeals to the device of question-and-answer. It may be truly said that this method as a means of clarifying Point of View (a term which is my own coinage) has been used by grammarians only sporadically; and no one, to my knowledge, has attempted to construct a set of questions which may be underlying predications.

Finally, I must state that I abandon entirely, at least at this initial stage, the conventional dichotomy according to which every predication may be divided into two parts: that which is known (psychological subject) and that which is unknown (psychological predicate). I have found it absolutely necessary to consider two other possibilities: none of the elements is known (e.g. *what happened?*); all of the elements are known (e.g. whether-or-not statements: see below).

1. What is the activity?

Z *What is done?* They made coffee, they took a vote; a tube was inserted, another bottle of wine was ordered, the matter is being considered; there was a scream of laughter, the telephone rang, somebody ran over a dog, several lines were omitted.

S *What does S do?* She bit her lips, she sent for the doctor, he threw open the clutch, I did some practising, he made fun of everybody, she teaches French, coffee stimulates, he circumvented the regulations, she slipped on the rug; so, I missed the bus.

O *What is done to O?* They put him in solitary, they gypped you; somebody kicked me; he was asked to wait, she was sent to a convent, he was applauded, we finally got waited on, he was overlooked; he was asphyxiated; it was folded in two, it's being painted; it got rained on.[7]

S, O *What does S do to O?*[8] I had it restored, I took them with me; the rest, he invested; I discouraged her; he depresses me, that cured me; I wrinkled it.[9]

---

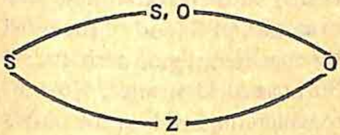
[6] I shall use 'subject' not only of the grammatical subject but also of the agent with passive verbs; and 'object' not only of the grammatical object but also of the subject of passive verbs.

[7] It should be obvious that O *(What happens to the object?)* is the natural habitat of the English passive auxiliary *get: it got rained on, we got waited on.*

[8] The four questions of group 1 represent, of course, only one question, with modifications of givenness *(what is done?)* and the answer to the four is basically the same: identification of activity. It should be noted, however, that 'activity' may be represented either by the verb alone, or by verb+object; with intransitives it is represented by the verb alone; with transitives often by verb+object – and always so, in Z and S. With the O-questions, however *(what is done, what does S do, to O?)*, it would seem to be the verb alone that describes activity: whereas in (S) *he scratched his head* the activity is that of head-scratching, in (O) *so I hit him*, the activity is that of hitting, not him-hitting. But these brief remarks do not begin to account for the problem: 'of what does the activity consist?' Still more tedious hair-splitting must be done before we can arrive at a clear, and perhaps ultimately simple, answer.

[9] These four questions could be doubled, and the three that follow could also be doubled, if we should take account of a fourth element that may be given: the adverbial element. One could posit not only Z, *what happened?* but also Za: *What happened there (here, the next day, when you said that)?* The presence of a given adverbial element may, in some cases, be of great importance.

What is the relationship between the four Points of View, which we have grouped under one heading (inquiry into activity)? Almost, we have an uninterrupted progression from the less given to the more: from zero given, to one element given, to two elements given. Almost, but not quite, for under 'one element given', there are two possibilities: S or O. Thus we cannot represent the progression linearly, for S and O must both be at the same distance from (the center of) Z. I see this as assuming the following form, with Z 'in the middle', extending around both sides through S and O, which meet in S, O:



At any rate, it is clear that we have a continuum of meaning, with one type melting into the adjacent one; since there is a single criterion and this criterion (givenness) is obviously relative, the twin facts of continuum and melting are accounted for.

Now, in the invented examples offered above, without context, I sought to choose predications as self-evident as possible, i.e. which would seem to be located in the central area of each of the four indefinite divisions: for the S- and O- types, I used mainly personal pronouns as subject or object, because of the high degree of givenness characteristic of such pronouns; for Z, the subject was either not mentioned or represented by an indefinite pronoun,[10] the object, too, being mainly indefinite. But it is one thing to invent examples supposed to illustrate a tendency; it is quite another to attempt to apply this tendency, unerringly, to a group of genuine utterances, as we find them; in the task of analysing 'real sentences', we must often hesitate: do *the baby began to cry*, *the baby was put to bed* represent Z: 'what happened, what was done?', or does the first announce the baby's reaction (S), the second, the treatment the baby received (O)? The context should contain the answer, which will hardly be so simple as to depend on whether 'the baby' had been recently mentioned. We must begin to learn to decipher context (including what follows as well as what precedes our predication); our reiterations of the importance of contextual analysis have been largely lip-worship; in practice, we have tended to leave this to the stylisticians.

It was said that, in each of the four groups, the examples were deliberately chosen with the view of making each group distinct from the others: this would seem to guarantee semantic homogeneity for each. Actually, however, I sought to introduce a rather high degree of heterogeneity: with S alone (and apart from the examples of imperfective aspect) we find: an involuntary activity *(she bit her lips)*, a 'step taken' *(she sent for the doctor)*, a stage in a process *(he threw open the clutch)*, an accidental activity *(she slipped on the rug)*, the results of an activity, accidental or deliberate *(so I missed the bus; he circumvented the regulations)*; and *I did some practising* offers a summary of how time was spent, while *he made fun of everybody* offers a summary of behavior. These examples belong together only because they all represent S; on other counts they must be separated. One day we may know how to break down all examples of S (and of the other types) into groups where they really belong together; then, perhaps, we will have made contact with Theme, and Purpose of Utterance. Now, let us turn to the second set of questions, those which inquire into the nature of the object.

---

[10] It may be said that when the subject is purely indefinite, an S-question is excluded: *somebody ran over a dog*. A problem may arise, however, when the subject is a loose *we* or *they*.

In order to know enough to ask about the object, the verb must already be given. This means that we have only two possibilities of givenness, not four, as in the first group: the verb alone is given (V); verb and subject are given (V,S):

2. What is the object?

V *What, who is —ed?* They served champagne at the dinner; in Spain they use a lot of garlic; (they took the vote and) John was elected; a new bridge is planned; Salesmen Wanted; a few scattered pages were found; in the margin was written: '–'; it is said that . . ., it was decided that . . . .
V,S *What does S —?* I see a mosquito, I want a match, I need a thimble, I'll take vanilla, I have some Energine here if you want it; and underneath it I found this clipping; I met Susie today; she reads nothing but trash; this year he's studying Ancient History.

It should be stated immediately that the formula '*what* is —ed, *what* does S —?' may be modified in one of three ways: '*which* (O) is —ed? *(they chose the last name on the list)*; '*how much* (O) is —ed, *how many* (O) are —ed?' ($1,000 *is being spent on it; it has two baths*); '*what kind of O* is —ed?' *(bad weather is forecast; she wears such old-fashioned hats)*.[11] The same modifications are possible in reference to human objects: *The younger man was hired, fifty guests were invited, professional players are excluded.*

If, in the first group above *(what is the activity?)* I sought to stress givenness by the choice of certain types of subject and object, here, where the verb is always given, I have chosen certain types of verbs. Just what type of verb it is that is most easily taken for granted (so that one is able, then, to put the object into relief as the point of information), I shall discuss elsewhere; in a following article I expect to offer a classification of the verbs characteristic of type 2. Even at this point, however, the reader will surely grant that the verbs illustrated are among those which seem to want to ask and answer our question: *what is it that is —ed?, what is the object?*[12]

---

[11] This would mean that in e.g. *She wears such old-fashioned hats* = 'the hats she wears are so old-fashioned', we have gained a new given element: the object itself; the unknown element inquired into is only the attribute of the object. Perhaps a separate set of questions should be set for this type.

But the factor of verbal material is not definitive, and we may often have to hesitate between *What is the object?* and *What is the activity?* Does *forceps were used in the delivery* tell what was used, or what was done (what step was taken)? Again, while *I ordered a milkshake* normally suggests the underlying question: *what did I order?* (what does S —?), the statement: *so, to kill time, I ordered a milkshake* (=‘and so I decided to order something: a milkshake, to be exact’) probably tells what I did. Nor, of course, within type 2 can one draw an absolute line between V and S, V; no more than between Z and S.[13]

If in the first general type *What is the activity?* we found four possibilities of givenness (Z, S, O; S, O), and in the second *(What, who is the object?)*, only two (V; S, V), in the third *(Who, what is the subject?)* we will find only one: in order to know enough to inquire into the identity of the subject we must already know the activity – which means, with transitive constructions, both verb and object.[14]

---

[12] It is, perhaps, too generally assumed that if an element is ‘given’, it is given by the immediate context; there is also the factor of human experience. Thus, with no preparation, I can ask ‘whom did you see today?’, when I could hardly ask, out of the blue sky, ‘whom did you choose?’ – Professor Spitzer tells me that one of the stock questions among faculty wives at provincial German universities used to be: ‘And what are you reading now?’ It was ‘given’ that a professor's wife must always be in the midst of a book.

Note, also, that certain type 2-questions are not apt to get asked under any circumstances: *Whom did you anticipate?, whom did you reassure?, what did he darken? what did he treat with care?* – except, of course, by one who is hard of hearing.

[13] In the abbreviated schema contained in this paper, I have made no special provisions for the combination of direct and indirect object; I should like to point out here, however, one of the possibilities that may occur in group 2: the indirect object being given, but not the direct. We would gain two new questions, parallel to the two already allowed for: in addition to V *(What was said?)* we could posit V, O *(What was said to him?)*; in addition to V, S *(What did you say?)*, V, S, O *(What did you say to him?)*. This would mean perfect symmetry between groups 1 and 2:

1. What is the activity?		*2. What is the object?*	
Z	What is done?	V	What, who is — ed?
S	What does S do?	S, V	What, whom does S —?
O	What is done to O?	V, O	What is —ed to O?
S, O	What does S do to O?	S, V, O	What does S — to O?

[14] To say that with transitive constructions, activity is represented by verb + object is not quite accurate (see footnote 7), though it happens to be true in this case.

3. Who, what is the subject?

V [O] *Who, what —s* [O]*?*[15] Mary found it, Mother made it; Don't look at me, John broke it; Hopkins won, Dean X is speaking, Marilyn Monroe is playing; (the door opened and) John appeared; it was built by the Romans, it was written by Shelley, the first steps were taken by UNESCO, it's signed by the secretary; he was raised by his aunt; he was driven to it by drink; it is produced by friction.[16]

Interestingly enough, this third type, in which two elements may be given, is very apt to melt into Z, where nothing is given: *the streets were guarded by Boy Scouts.* If we may take for granted the situation of guarded streets, this answers V,O; if not, Z.

With this, we have collected seven questions: one in which three elements may be unknown (Z), three in which two elements may be unknown (S, O, V), and three in which only one element is unknown (S, O; S, V; V [O]). This is as far as we can go in accounting for predications in which at least one of the three elements, S, V, O, is unknown.

But it is not said that one of the three elements must always be unknown; it is quite possible to find predications in which all three are taken for granted, but which nevertheless offer new information. Such statements fall into two main groups (a) the informational element is represented by an adverbial expression: *I bought it at Wanamaker's, I bought it yesterday, I bought it on the installment plan, I bought it because I need it, I bought it with Mother*; (b) the information consists in predicating 'true or false': *No, I didn't see him, I won't take it, I don't believe so; I'll surely do it, I believe you, I'll find it!* For these two types, I shall not use symbols (V, S, O would have to do duty for both): I shall speak of the first as 'adverbial variants', of the second as 'whether-or-not variants'.

And I call them variants, i.e. variants of the seven basic questions just considered, because I believe that it may be possible, though the task is often difficult, to assign any given adverbial or W-N type to one of the basic questions. For example, if *I took a bath* tells what I did, then *I took my bath in the kitchen* would be an adverbial variant of S: though the immediately underlying question is *where did I take my bath?*, underlying this, in turn, is (the already answered) *what did I do?* But with *he was operated on in the kitchen, imagine!*, we would have an adverbial variant of O; underlying *where was he operated on?* is the basic question (already answered) *what happened to O?* And *I found it in the garden* would be an adverbial variant of (the answer to) *what did I find?* – whereas *I didn't find it* would be the corresponding W-N variant; *oh yes, he was operated on*, would be a W-N variant of O, and *I've already taken my bath (I haven't taken my bath yet)* would be W-N variants of S.

---

[15] I bracket the O of V, O, because with intransitive predications only the verb would be given.

It will have been seen that of the six preceding questions, only two apply to intransitives: Z and S. Thus there exist, in all, only three basic questions that can be answered by intransitives: *What happens? What does S do? Who (what) —?*

[16] When the subject is identified, just as when the object is identified, we may count with the threefold modifications: *which, how many, what kind of . . .?*: e.g. *The ugliest one got married first; fifty people answered the ad.; a very stupid clerk waited on us.* (It may be that the same modifications will apply when the activity is identified.)

If I am right in conceiving these two types as variants, and if it is possible for each of the seven basic questions to have an adverbial and W-N variant (the two hypotheses are not the same), this would mean that we have arrived at a total of twenty-one questions. Actually, we have only begun the process of multiplication: in the first place, there are two types of W-N questions: the straight interrogative and the indefinite interrogative; I may ask, for S, not only *Did she do this?* but *did she do anything?*; for V,S not only *did you find it?* but *did you find anything?* for O, not only *was this done to him?* but *was anything done to him?* etc. Thus, for example: *I was awfully tired but I did manage to scrape the dishes before I left* answers neither the basic question *what did I do?* nor the straight interrogative W-N: *Did I scrape the dishes?*, but rather the indefinite W-N: *Did I do anything (tired as I was)?*[17] Secondly, it is possible for W-N to offer variants not only for the seven basic questions, but also variants (and both kinds of variants) for the adverbial variants! Consider, for example, the basic question (V, S) *Whom did you see?*, of which an adverbial variant would be *when did you see him?* Converted into W-N this may appear either as *Was it on Tuesday that you saw him?* or as *Did you see him at any time (ever)?* This would give us, theoretically, a total of forty-two questions.

If the outline I have offered is to be taken seriously as theoretically able to account, at least in its own terms, for the variety of utterances in any language, it is obvious that it stands in need of many modifications; in my presentation I have completely neglected (i) constructions with copula (none of our questions can apply, without re-wording, to *it is pink*), and (ii) dependent clauses or imperative sentences; nor have I considered cases of compound underlying questions. Moreover, even within the types of statements chosen, the examples offered have suggested only a tiny fraction of the variety and complexity that are to be met with, once we decide to dive into the ocean of What Is There (and the ocean of predications *is* language). But I hope to have shown, at least, that it is possible to envisage for the future the classification of predications, systematically, in terms of meaning (it is, of course, only by appeal to meaning that we can ever become systematic in syntax): a classification by means of which one unit, in its variety, is continuously measured, in which one criterion (at least) is forever relevant – by which, in other words, we are assured of a continuum.

---

[17] With this type of W-N variant *(Did I do anything? – Yes, I did manage to scrape the dishes)*, it would no longer be true that all elements are given.

This schema has been presented primarily from a semantic point of view: I have sought to distinguish between various types of meaning, with little or no concern for form. But if these distinctions are to work for syntax, they should be applicable somehow, to some degree, to the already existing, familiar, forms of language: our so-called constructions.[18] Whatever the degree of applicability, the reader is surely aware of the apparent paradox involved in the attempt to apply the meaning of a sentence to the form of a non-sentence (a paradox made necessary by the fact that only with the sentence do we find a continuum of meaning, only with constructions do we find easily recognizable, simple forms); and it is obvious that not one of the forty-two possible distinctions will be found to correspond directly to a given construction. Just how a correspondence might be established between constructional form and sentence-meaning has already been suggested: a given construction would be described in terms of the types of sentences in which it may appear: in terms of the types of sentence-meanings which it may help to express. This sum total would be the (rough) meaning of the construction.

As for the degree of its applicability, I have already said that I believe it will work for constructions of nexus, and I hope it may work for those of junction; I may add that I know it will work for two types of nexus problems, to the analysis of which I have devoted myself for the last few years: patterns of word order involving position of subject and object,[19] and constructions of the 'indefinite human agent' (the passive; French *on*, Spanish *se*, English *you*, etc. etc.); it was during the course of this twofold investigation that the schema offered above was developed. (Thus the system presented here from a purely semantic point of view is actually the result of an inquiry into Form: many different forms at once, in several languages.) I shall choose the first of these two constructions to illustrate, very briefly, certain possibilities of correspondence between sentence-meaning and constructional form.

In English, for example, the inverted object may appear (apart from the variants) in group 3 *(that one, Mother made)*, in group 2 *(A fine chance you have)*, and in group 1 *(the rest, I invested)*; in groups 3 and 2 all possibilities are represented, but in group 1, it seems that only S, O appears: compare also *the last one, I kept*; *the lining, I had dyed*; *the furniture, I sold*; *the first chapters, I re-wrote*; *the dog, we're taking with us*; *the inside piece, I broke*; I cannot easily imagine O (no subject given: **The furniture, the insurance people confiscated*), and even less, S or Z. Again, I have found that in Spanish VOS answers none of the group 2 questions, but appears in group 1 (*Aquí dió un salto 'El Nacional'*: at this, gave a start 'El N') and group 3 (*Dirige las maniobras el general Hoge* 'directs the manouvres Gen. H'); VSO, which appears in group 1 (*Cogió Juan su bastón* 'Took J. his cane')[20] and group 2 (*Viste el rey traje azul* 'wears the King blue apparel'), is apparently excluded from group 3; both, in answering group 1 questions, are limited to S and Z: neither is found with O or S,O.

---

[18] I would say that, even if no correspondence between meaning and form could be shown to exist, the classification of sentence-meanings here proposed would still not be without value: to ask ourselves (finally!) what a sentence means, might be a good logical exercise in itself; with more experience in applying it, we might succeed better in teaching our children how to read.

[19] In referring to word order, I use 'subject' and 'object' in their conventional meaning.

To a certain extent, then, the system serves to distinguish, definitively, between VSO and VOS in Spanish (and in English, between OSV and VSO); in other cases, the criterion of Point of View is apparently useless: both VSO and VOS being found in S and in Z. But we may be sure that when the two constructions alternate within the same sub-division, each will have its own way of answering the particular question; and the final description of the construction would include not only the list of questions answered but the particular way in which these are answered. If this 'way' can be analysed (an analysis involving criteria other than Point of View, but greatly facilitated by the breakdown afforded by Point of View), then we can hope to contrast VSO and VOS even within the divisions where they alternate. Thus our system, first directly, then indirectly, would have thrown light on constructional usage; but it is just as true that the constructions will have thrown light on the system: the at present completely vague 'What happens, what does S do?' will become more concrete, more objective, as we come to know by the study of more and more constructions, all that such questions may mean for a given language: we can cut a huge chunk, labelled S or Z, out of the language (a chunk of predications, grouped according to constructions) and examine it to our heart's content, from every possible angle. It has never been possible before to cut a piece out of the language and look at it.

---

[20] That in *Cogió Juan su bastón*, the word order may have direct relevance to our questions, the reader will surely grant as a possibility. But what, for example, of the possessive adjective *(su bastón)* which also happens to be found in the predication: does this, too, play a discernible role in this sentence – considered as an answer to an S-question?

In a sense, it does. I have already shown (Hatcher, 1956, p. 6) that, with OVS, the rare cases of possessive adjective + object have a very special semantic range, which excludes the idea of mere possession (as in *su bastón*): I have shown, in my monograph, that OVS does not answer S-questions. Here, perhaps, we have an illustration of what was discussed as a possibility in footnote 4: the ultimate relevancy of sentence-meaning for problems of junction.

One may consider also the conventional distinction between definite and indefinite words: this must correspond, roughly and partially, to the distinctions of our questions, since the criterion, in both cases, is that of givenness. One day it may be possible to state precisely in which variations of which questions the subject or object may be definite or indefinite.

All this is predicated, of course, upon our ability to recognize the different semantic sentence-types when we see them. This will not be easy: 'real sentences' do not come in little clusters of obvious types such as I offered above. Mainly, we mentalists do not understand our sentences (I speak from a vast experience of not understanding); mainly, we do not know yet how to find the clues that the patient context is always offering us. But we can learn: learn new patterns of association, learn to think, at first, in more fluid categories – which will gradually assume clear, firm contours.[21] 'Our subjective must put on objectivity'.[22]

The objectivity flaunted naïvely or impudently by those who work with ready-made, fixed categories, is an objectivity with which the scholar begins, with which he has already begun before he begins to think: it replaces for him the necessity of thinking. The objectivity to which we may one day hope to attain is the objectivity with which the scholar ends – or, goes forward. Only that objectivity which is a result is to be cherished: one must always begin by being subjective, that is, creative. The rest is sterility, the curse today of Syntax, which is gradually assuming the nightmarish aspect of a Waste Land.

By ready-made categories I mean not only the antiquated distinctions of conventional syntax used by the statisticians, but also the newer, stylish,

[21] We may even come to know what we mean by 'givenness', about which there has been so much talk. It has already been suggested, as obvious, that we must expect givenness to be a matter of degree; to this it was added (footnote 12) that an element may be given by experience as well as by the context (the two are not mutually exclusive). But it is also true that what seems to be logically given may not coincide with what is given for the speaker: not only may he jump to something which he decides, privately, to take for granted; it may also happen that the obviously given, even when mentioned by him, is only casually given for him: e.g. *he looked at me and gasped*. I should say that this, like *he looked up and gasped*, tells only what S did, not what he did to me (or even, in regard to me): here, though *me* is given, it is not given as Point of View. And if it comes to a clash between givenness and Point of View, the latter must win out: this is absolutely necessary, so long as we have recourse to questions.

[22] It should be clear that I am thinking always in terms of a gradual development of our senses of discrimination. We learn from one construction to sense a certain distinction and nothing more; with this we may go to a second, which may teach us a new distinction, which, in turn, may enable us to return to the first with slightly keener vision. It's all there for us to see; we must simply practice at training our vision. We are like new-born babies who have not learned yet how to focus, and who, for practical purposes, are blind. The baby 'sees' first those objects with the most obvious outlines and most *voyant* qualities – comparable to what the statistician sees in syntax. The crime of the statistical method is that it condemns the baby to see forever only huge red and yellow blocks.

imported categories which some scholars, desperately or frivolously, and with varying degrees of understanding, have sought to borrow for Syntax from other fields where they were worked out in accord with principles relevant to those fields. Such principles cannot be relevant to the real problems of syntax and they produce nothing new except new jargon. Waste Land and Babel!

Syntax is no ancillary science that can only borrow and attempt to apply the results of others' thinking (whether coming from phonemics, or the elegant field of Symbolic Logic, or the less elegant field of Gallup Poll sampling). Syntax is, or must become, a basic science: a science in the Humanities, from which the other linguistic disciplines (and perhaps also Psychology and Humanistic Logic), enjoying, as they do, less direct contact with meaning that reveals itself systematically in concrete forms, will one day be glad to learn.

References

GOUGENHEIM, G. (1938), *Système grammaire de la langue francaise*, Paris.

HARMER, L. C. (1954), *The French Language Today*, Hutchinson.

HATCHER, A. G. (1956), 'Theme and underlying question', *Word*, vol. 12, supplement.

LE BIDOIS, R. (1952), *L'Inversion du sujet dans la prose contemporaine*, Paris.

4 V. Waterhouse

Independent and Dependent Sentences

V. Waterhouse, 'Independent and dependent sentences', *International Journal of American Linguistics*, vol. 29, 1963, pp. 45–54.

Introduction

In any language, some structures are INDEPENDENT; others are DEPENDENT. Both kinds of structure are recognized on various levels of language. Within words, free morphemes *(boy, come)* are independent; bound morphemes *(un-, -couth, -ed)* are dependent. Within certain phrases *(big John)*, the head *(John)* is independent; modifiers *(big)* are dependent. Within clauses, the predicate nucleus *(come, run)* or the subject-predicate combination *(John came, Birds sing)* is independent; optional peripheral expansions such as expressions of time *(at night)*, or location *(in town)*, are dependent. Within sentences, there are both independent and dependent clauses. The concept of dependence for these levels is recognized by most scholars.

Where sentences are concerned, however, the Bloomfield definition that 'Each sentence is an independent linguistic form, not included by virtue of any grammatical construction in any larger linguistic form', seems to have the status of a law of the Medes and the Persians (1933, p. 170). This has resulted in an atomistic preoccupation with units no larger than the sentence by scholars from Bloomfield all the way to the exponents of transform grammar. Recent analysis in some aboriginal languages, however, has made it obvious that not all sentences are independent (Loriot and Hollenbach, 1970), but that there are in fact grammatical constructions which indicate that some sentences are dependent and therefore to be included in some larger linguistic form such as paragraph or discourse.

Types of dependent sentences

As a matter of fact, Bloomfield himself mentioned certain kinds of structures or items which can now be seen to mark certain sentences as having a dependent relationship within a larger unit. One of these structures is called by Bloomfield a completive sentence type: 'The completive type consists of a form which merely supplements a situation – that is, an earlier speech, or gesture, or the mere presence of an object. . . .' (1933, p. 176). A second item is the anaphoric or dependent substitute, one which implies an antecedent, or 'a recently uttered form' (1933, p. 249). A third item is a sequence-marking particle or phrase, such as *therefore, moreover, consequently, on the other hand,* and the like.

On the basis of these structural items (taken in reverse order), dependent sentences can be recognized as consisting of three formal types:

1. ADDITIVE, formally marked by a sequence-marking particle or phrase;
2. SUBSTITUTIONAL, formally marked by the presence of an anaphoric substitute without identified referent in the same sentence (or by absence of subject or object noun in languages which do not employ such substitutes);
3. TRUNCATED, marked by an incomplete structure, i.e. composed of only part of the structure considered obligatory to a full independent sentence.

Functionally, additive sentences may be termed SEQUENTIAL; substitutional sentences, REFERENTIAL; truncated sentences, COMPLETIVE. Combinations of types also occur in some languages. In the balance of the paper, the functional labels will be used to designate dependent sentences.

Just as languages differ in the structural patterns of their morphemes, words, phrases and clauses, they differ also in the structural patterns of both independent and dependent sentences. There are differences in types of sentences found from one language to another, and in the frequency of use of one type rather than another. Also, within the same language, style and type of discourse may govern the use and frequency of one type in preference to another.

For example, all cited types of dependent sentences are found in English; completive sentences are common in Campa of Peru but do not occur in Tzeltal of Mexico; sequential sentences abound in Greek, and in Zoque and Aztec of Mexico. In English, completive sentences are common in conversation and in popular narrative style, but rare in formal discourse or in textbook style, where sequential sentences are more likely to occur. English sequential sentences differ radically from Greek ones in terms of frequency, however. Most sentences after the first in a Greek text must contain some sort of connective, whereas in English such widespread use of connectives or sequence markers is not considered good style, and the use of *therefore, moreover, however* and *on the other hand* marks special summation, contrast, transition, emphasis, and the like. Referential sentences do not appear to be restricted to any particular style or universe of discourse.

Both sequential and referential sentences are recognized by linguists and native speakers as full, structurally complete sentences. Completive sentences are recognized either (a) by native speakers as contextually conditioned sentences (such as answers to questions); or (b) by analysts as phonological sentences – with the same intonational and junctural features as those found with structurally complete sentences.

Dependent sentences in English

In English (and possibly in other languages as well), it is necessary to recognize the distinction between INDEPENDENT versus DEPENDENT on the one hand, and COMPLETE versus INCOMPLETE on the other. Both independent and dependent sentences may be structurally complete or structurally incomplete. Independent sentences are defined distributionally as those which can occur as a complete utterance without ambiguity, or which can initiate discourse without necessity for some type of defining context. Dependent sentences are defined as those which cannot occur as a complete utterance or initiate discourse without some defining context, and which are formally marked for dependence in one of the above ways.

Independent complete sentences in English are of the typical NP+VP type *(The Phoenicians were a sea-faring people, Jack threw the ball straight at Susie)*. Independent incomplete sentences are (a) those in which the subject pronoun (or pronoun and auxiliary verb), or the first two or three unstressed syllables are omitted *(Just got back from Austin yesterday, Gonna be hot today, Coming?)*; (b) aphorisms *(The more the merrier, Never a day without a change)*; (c) exclamations *(Fire! Here!)*.

Dependent complete sentences in English are sequential *(Consequently, there is no water left in the well)*, referential *(He is every inch a man)*, or sequential referential *(But they were gone)*. Dependent incomplete sentences are completive *(Very nice, really)*, sequential completive *(But not now)*, referential completive *(Not he)*, or sequential referential completive *(But always her own)*.

Complete sentences, whether independent or dependent, are more common in formal discourse than are incomplete ones. Incomplete sentences, both independent and dependent, are characteristic of informal discourse, both written and spoken.

Some sentences may be said to be contextually dependent, in that they are not easily understood except in the context of the larger unit (linguistic or non-linguistic) in which they occur. Unless they can be said to be unambiguously marked for dependency in one of the ways mentioned, however, they are not considered to be within the scope of the present discussion. Some contextually dependent sentences are marked by demonstratives or the definite article (Fries, 1952, pp. 241–52), but not all are so marked. Furthermore, occurrence of demonstratives or of the article does not always clearly mark dependence.[1] Such sentences should ultimately be analysed within some framework, but they are omitted from further discussion here.

---

[1] Demonstrative pronouns perhaps mark dependence more often than do demonstrative adjectives. Similarly, first and second person plural pronouns may sometimes

For analysis of dependent sentences in English, I have employed a corpus of random utterances, culled from various books and periodicals. While it is obviously possible for a native speaker of a language to use himself as his own informant, it has been my observation that sentences which native speakers of English who are also linguistic scholars frequently offer as examples of English structure sometimes bear little resemblance to the sentences spoken by the average man in the street. Therefore, I have chosen for the most part to present examples of English utterances found in written material of various types.[2]

Dependent sentences in English include the three major types mentioned: sequential, referential, and completive, plus combinations of these, marking combined types sequential referential, sequential completive, referential completive, and sequential referential completive.

I Sequential sentences in English are introduced by a wide variety of sequence markers.[3] In the following examples, the sequence marker is italicized. In some examples (especially with *however* and *too*), the sequence marker does not introduce the sentence. In a few other examples (with *thus* and *nor*), the normal word-order is changed. In a few examples, some words have been omitted in the interests of brevity; the omission does not affect the status of the sentence in any way. Examples are numbered, and the letter S (for sequential) is preposed. In cases where greater clarity is considered necessary, the preceding sentence from the context is included in parentheses.

S1. *Hence*, we have no recourse but to accept language as a fully formed functional system. . . .

S2. *And so*, the process has continued uninterruptedly down to the present day.

---

signal dependence, as in such utterances as 'Mother, Aunt Elsie and I went down town. We enjoyed it!'

[2] The corpus used includes such diverse items as: Edward Sapir, *Language*; C. C. Fries, *The Structure of English*; E. A. Nida, *Synopsis of English Syntax*; R. E. Pittenger, C. F. Hockett, and J. J. Danehy, *The First Five Minutes*; William Strunk, Jr and E. B. White, *The Elements of Style*; *The Best Short Stories of 1941*; Sinclair Lewis, *Arrowsmith*; Mary Roberts Rinehart, *The Breaking Point*; Edith Wharton, *The Greater Inclination*; J. F. Cooper, *The Last of the Mohicans*; Stan Delaplane, editorial in the *Oklahoma City Times* for 4 July 1960; and numerous issues of *Reader's Digest*.

[3] The mere presence in a sentence of a sequence marker does not thereby constitute the sentence a sequential type – it may signal a relationship between clauses or phrases within the sentence (I thought I'd go *but* I changed my mind). But in most cases it is obvious when the sequence marker signals a connection with the previous sentence (as in this one).

S3. *Similarly*, the eight majority traits are divisible into two subgroups.
S4. One can, *however*, compare languages which cannot be shown to be related.
S5. (We do not sell fireworks in our enlightened community.) *Therefore*, we have to drive over to another county where the inhabitants are crude and full of fun.
S6. *Besides*, I was already a success, and rather enjoyed the illusion.
S7. *For* a vast military disaster was in the making.
S8. *And* the Germans encouraged this lightheartedness.
S9. *But* snail fever responds slowly to treatment.
S10. *In addition*, my personal ties to you are close.
S11. *On the other hand*, a rising lower middle class of city dwellers . . . are becoming aware of what they have been missing. . . .
S12. *As a result*, resentful Latin Americans accuse North Americans of exploitation. . . .
S13. *Thus* did a startling merchandising explosion burst across the country in the fall of 1959.
S14. *Also*, the very poverty of the raw-material countries threatens the prosperity of the rich.
S15. *Nor* was Ma above doing a little promotion.
S16. *Yet* for some unfathomable reason, this duck chose to nest in the roaring heart of downtown Milwaukee.
S17. There were dogs in the crowd, *too*. . . .

II Referential sentences in English are marked by anaphoric substitutes of two types: (a) third person pronouns; (b) pro-verbs. Sentences containing pro-verbs may be analysed in one of two ways: as referential sentences, or as completive sentences. The choice of analysis depends on whether one considers the pro-verb to substitute for the verb in the same fashion as the pronoun substitutes for a noun, or whether one considers the occurrence of the pro-verb to constitute structural incompleteness. There is probably difference of opinion among native speakers of English on this point, but the referential analysis is here chosen as preferable.[4]

In sentences marked by pronouns, the pronoun may be subject, object or possessive. More than one pronoun may occur in the same referential sentence; they may refer to the same antecedent or to different ones.

---

4. I am indebted to my colleagues Barbara Erickson and Doris Bartholomew of the Summer Institute of Linguistics for the suggestion of this analysis. I had originally set up the pro-verb sentence as completive, but found the analogy of the pronoun and the pro-verb very appealing. The further note from Bloomfield (1933, pp. 251 ff.) considering pro-verbs as anaphoric substitutes tipped the scales in favor of the referential analysis.

Referential examples with pronoun are indicated by RN; each pronoun is italicized.

RN1. *He* looks something like a near-sighted kipper.
RN2. *She* didn't complain.
RN3. *It* has had a massive influence.
RN4. *They* sprawl, rather than sit.
RN5. Word spread that *he* was being shoved aside.
RN6. In the morning *they* had all gone.
RN7. Perhaps *they* are not yet prepared to inherit the mess.
RN8. With each passing year, *he* became more cantankerous, *his* beard more scraggly, *his* clothes more rumpled.
RN9. *Her* dream came true first.
RN10. *Their* policy is to help a boy at least once. . . .
RN11. There, death claimed *him* in 1957. . . .
RN12. The chains of *her* lifelong enslavement fell from *her*.
RN13. Some of *it* will never be understood.
RN14. We must not risk weakening *them*.
RN15. No cars pass *it*.
RN16. Always *they* could make *themselves* understood in *their* writing.
RN17. Chief among *them* was the absolute prohibition of marriage.
RN18. *He* tries to scare *him* off.

Sentences marked by pro-verb are less frequent in the corpus than those marked by pronoun. Examples are indicated by RV; pro-verbs are italicized.

RV1. I honestly *would*.
RV2. We all *do*.
RV3. Now I *do* not.
RV4. Well, perhaps you *will*.
RV5. Many *are* not.
RV6. If you *don't*, someone else *will*.
RV7. Bob probably *should*.

Some sentences are marked by both pronoun and pro-verb. These may be called double referential; examples are indicated by RR. Both pronoun and pro-verb are italicized.

RR1. *They would* and *they did*.
RR2. *They* really *did*.
RR3. Yes, *he will*.
RR4. Well, *he doesn't*.

III Completive sentences in English may consist of any part of an independent sentence which can constitute a complete unit on a lower level. Most examples found in the corpus are of phrases or single words; one, however, is a relative clause: *Which by no means meant that it wasn't sharp or progressive.* Completive sentences occur in two characteristic distributions: (a) as responses; (b) as expansion of some part of the previous sentence.

Completive sentences as responses are usually in answer to questions, but may be responses to comments, or even the continuation of a previous conversation. In the following examples, completive sentences are indicated by C, and the preceding question, comment, or context is enclosed in parentheses.

C1. (Which one was it?) *This one.*
C2. (What about the baby at the Homer place?) *Not due until Saturday.*
C3. (How does it look?) *Pretty good.*
C4. (Who came in?) *The kids next door.*
C5. (What shall we have for dessert?) *Cherry pie.*
C6. (Don't you think we should call the doctor?) *Not yet.*
C7. (Is there anything important coming up?) *No, not at all.*
C8. (Can you bring it over?) *Yes, right away.*
C9. (He says it'll take two spools.) *Just about, I reckon.*
C10. ([Previous discussion of a not very neat room.]) *Neat as a pin this morning.*

Completive sentences as expansions have been observed to expand Subject, Predicate Nominative, Object, a Descriptive Item, Time, Location, and, rarely, Predicate. Expansion processes include addition, substitution, repetition, negation. Not all processes are exemplified for all types. Processes may be combined; repetition and addition frequently occur together. Examples of expansion types are numbered by decade classes 20–80 respectively.

Class C20 are expansions of Subject.

C21–3 are examples of addition.

C21. (Mr Lawyer, you knew he didn't pay me.) *A man with your education.*
C22. (The rest was nerve-gnawing.) *Droning afternoons of hospital demonstrations among stumbling students barked at by tired clinical professors. . . .*
C23. (The fence itself wasn't much . . . but behind the fence were deeds in the county courthouse. . . .) *Property rights.*

C24–6 are examples of addition and repetition.

C24. (Was it possible that these were his friends?) *These mincing women, all paint and dye and whalebone, these apathetic men who looked as much alike as the figures that children cut out of a folded sheet of paper?*
C25. (Well, that's the thing you remember.) *That, and the sprinkling carts, and the brown afternoon in the street, and the old woman who sold hot chestnuts. . . .*
C26. (Something had happened, she was sure.) *Something connected with Dick.*

C27 is an example of expansion by negation and repetition.

C27. (The letter was the only thing in it.) *No powder, no money, no lipstick, nothing but the letter.*

Class C30 are expansions of Predicate Nominative.

C31 is an example of addition; C32, of substitution; C33, of repetition; C34–5, of repetition and addition.

C31. (That was a fine place.) *Mostly garden.*
C32. (In the Bronx where Millie Patron grew up, a girl who wanted to work knew she would be either a school teacher or a stenographer.) *One or the other.*
C33. (Mother said it was a nice place.) *A real nice place.*
C34. (He didn't want dreams; what he wanted was facts.) *Facts and hope.*
C35. (I'm telling you it's a duck.) *A mallard duck with three eggs already.*

Class C40 are expansions of Object.
C41–4 are examples of addition.

C41. (He told himself that he didn't want much.) *Only peace, work, and a place.*
C42. (She cut them, sewed them, finished them in various sizes and patterns.) *Soft collars to be attached to men's shirts.*
C43. (His eye fell on 'Gertie's window' in Gimbel's store.) *Just the thing!*
C44. (He came to know all of Leora's background.) *The hamlet of Wheatsylvania, North Dakota; one street of shanties with the red grain elevator at one end.*

C45–7 are examples of repetition and addition.

C45. (He sniffed at the white mist and smelled that exhilarating friendly odor which morning distills from a peasant village.) *An odor compounded of wood-fire smoke, cattle, and a mossy smell coming from the damp roofs of shingle and thatch.*

C46. (It would be good to be on the loose in a town like Paris with a uniform on and no worry at all.) *No worry about money that is.*
C47. (The town, which had seemed inimical for so long appeared to him suddenly to be holding out friendly hands.) *More than friendly hands. Loving, tender hands, offering service and affection and old time friendship.*

Class C50 are expansions of some Descriptive Item or of the sentence in general.

C51–3 are examples of addition.

C51. (The Fourth of July these dependent days simply reminds us how much independence has been blown out of our youth.) *Surrendered to safety, sanity, and the supermarket.*
C52. (Life went on, the comfortable life of an upper middle class household.) *Candles and flowers on the table and a neat waitress to serve. . . .*
C53. (They had worked to get the crowd into just the right mood for her.) *Laughter, then the big scare.*

C54–5 are examples of repetition and addition.

C54. (He sat on the steps of the courthouse in Lilicutum and read *Possession* and thought it would be wonderful to know the author, but impossible.) *Just impossible.*
C55. (They were small, dark men, polite.) *Polite till they knifed you.*

C56 is an example of negation followed by addition.

C56. (The bull seemed innocent enough. . . . His white coat gleamed like thistledown.) *Not a scratch on him, and probably not a staple jerked out of the fence. Over clean as a greyhound with those long jumping legs.*

A special type of expansion of description is by the addition of a comment a little more remotely related to the specific content of the preceding sentence than the above examples. C57–9 are of this type.

C57. (Very soon they would be alone together again, as in the early days before the children came.) *All the years and the struggle, and then back where they started.*
C58. (Judson Clark had been famous, notorious, infamous and dead, all in less than two years.) *A shameful and somehow a pitiful story.*
C59. (Put the other way, the sentence becomes stiff, needlessly formal.) *A matter of ear.*
Class C60 are expansions of Time.

C61 is an example of addition; C62, of repetition; C63, of negation.
C61. (I'd like to see you again.) *Tomorrow night maybe.*
C62. (Like one who enters a room for the first time, to find it already

familiar, for a moment he felt that the thing he was doing he had done before.) *Only for a moment.*
C63. (Take garden seeds for example. I can remember when all you had to do was stick them in the ground.) *Not any more.*

Class C70 are expansions of Location.
Examples are of addition or substitution.

C71. (It's near there he lives.) *Around the corner, in fact.*
C72. (They had come back, never to be lost again.) *Home.*

Class C80 are expansions of Predicate.
Only two examples occur in the corpus; both are by negation.

C81. (You'd think the salesmen would be working every possible minute.) *Not a bit of it.*
C82. (Appalled, the lawyer tries to buy the brute off.) *Nothing doing.*

Occasionally a series of completive sentences occurs. In the following example, each sentence is an expansion of the one immediately preceding, by the following processes: (a) of time by addition to the negative time particle; (b) of predicate by repetition and addition; and (c) of location by addition: (I'll never let her lose anything.) (a) *Not until I can take her away.* (b) *Take her away from the circus.* (c) *From the air.* Note also examples above.

Completive sentences may constitute an entire conversation,[5] as Nida exemplifies (1960, p. 216): *How much for these? 50¢ a dozen. Too much. How about these over here? Well, how much for them? 40¢ per. All right. A few sprigs of parsley too then? Okay. Thanks. Goodbye.*

IV Sequential referential sentences combine sequence markers with referential substitutes. Examples are indicated by SR. Both sequence marker and substitute are italicized.

SR1. *But it* was behind the sea wall.
SR2. *And* one year *she forgot.*
SR3. *They too* are on vacation.
SR4. *It* has some astonishing effects, *however.*
SR5. *So* this is a sort of public acknowledgement of my obligation to *him.*
SR6. *It therefore* wastes no time talking down to adults. . . .
SR7. *But she* had obviously come to protect us from *him.*
SR8. *It* made a great many other people feel good, *too.*

---

[5] An admirable example is a drama by Harold Whitehall of Indiana University in which all speeches consist of single words, almost all of which are completive sentences. Regrettably, it is a little long to reproduce here.

SR9. *And* we told *her* fine.
SR10. *But* everybody *was.*
SR11. Mine *don't* really, *though.*
SR12. *Besides* Bill *can't.*
SR13. *Furthermore, they* obviously *should.*
SR14. *So did* just about everyone else.
SR15. *And,* believe me, I *was.*

V Sequential completive sentences combine sequence markers with incomplete structure. Examples are indicated by SC; the preceding context where necessary is enclosed in parentheses.

SC1. (. . . he had had an idea all along that his mind was running too fast for him, sort of, so that he was overlooking things.) *And maybe important things.*

SC2. (A literal translation of this sheet of water . . . would be 'The Tail of the Lake'. Lake George, as it is . . . now . . . called, forms a sort of tail to Lake Champlain, when viewed on the map.) *Hence the name.*

SC3. (And then, like as not, she'd get up, slam the door to, and go off to see one of her girl friends or something.) *But not now.*

SC4. (If you know how to do the right things.) *Or if you know the right people.*

SC5. (That was one thing he mustn't do; without even thinking about it, he knew he mustn't run.) *Or talk.*

Both completive and sequential completive sentences may occur in sequence: (We pinched [the firecracker] and lit it with a piece of Chinese punk. CRACK! There you stood.) *Still shivering. But a hero.*

Sequential completive sentences may be used to resume a narrative after a digression:

SC6. (Is the fusing technique thereby set off as the essence of inflection? I am afraid we have not yet reached our goal. . . . [Followed by five paragraphs of discussion of agglutination.]) New paragraph (with sequence marked by the verb): *To return to inflection . . .*

SC7. (I was just leaving Oliphant and Company's offices when I saw Mike . . . [Followed by seven sentences of digression.]) New paragraph: *But about Mike. . . .*

Sequential completive sentences may also occur in paragraph initial to compare or contrast a new paragraph with the material of the preceding paragraph:

SC8. (In spite of the fact that the greatest race cleavage in this region lies between the Papuans and Polynesians, the major linguistic division is of

Malayan on the one side, Melanesian and Polynesian on the other.) New paragraph: *As with race, so with culture. . . .*
SC9. (To put it concisely, walking is an inherent, biological function of man.) New paragraph: *Not so language. . . .*

Occasionally, both sequential completive and sequential types may be combined into a single complex dependent sentence: *At least to begin with, for habit soon makes the association nearly as automatic as any and more rapid than most.* It may be noted that the above example is contextually dependent as well as structurally so, for only context gives a clue to what is being discussed, yet there seems to be nothing in the structure of the underlying sentence (*Habit makes the association automatic*) which can be said to mark a dependent relationship.

VI Referential completive sentences combine referential substitutes with incomplete structure. Examples of these were not found in the corpus examined, but the following are illustrative. Examples are indicated by RC; substitutes are italicized.

RC1. Not *he.*
RC2. Forever *theirs.*
RC3. In *her* own house.
RC4. For *him* as long as *he* would stay.
RC5. Sure *did.*
RC6. Never *will.*

VII Sequential referential completive sentences combine sequence markers, referential substitutes, and incomplete structure. These also were not found in the corpus, but could undoubtedly be found with further search. Examples are indicated by SRC; sequence markers and substitutes are italicized.

SRC1. *But* never *them.*
SRC2. *Their* own car, *however.*
SRC3. *And his* own uncle at that.
SRC4. *Her too.*
SRC5. *Didn't either.*
SRC6. *But* sure *was.*

Dependent sentences in other languages

We now turn to examples of dependent sentences from languages other than English. Examples are cited only by English gloss (with gloss of preceding sentence given in parentheses where needed), since native form is not crucial to the discussion.

For Campa of Peru, in which the concept of dependent sentences first came into current focus, four types of dependent sentences are described (Kindberg, 1961, pp. 505–18). Emphasis, Result, Interrogation and Response. Emphasis and Response contain completive sentences; Result is sequential, marked by a special form of the verb; Interrogation is functionally dependent, since it requires a response. The Emphasis type consists of some item from the previous sentence which is being emphasized: *(He cultivated-it, yucca, corn.) Everything*. The Response type is in answer to a question: *(Where are you going?) Over there*. The Result type presents the result of action described in the previous sentence: *(He blew on it a little and it became metal.) As a result it was hard*. The interrogation type consists of information queries: *What is wrong with you?* For specific Campa data, the reader is referred to the article cited.

For Chontal of Oaxaca, Mexico, two classes of dependent sentences have been described, both completive (Waterhouse, 1962). The Response type may consist of a single word or phrase; occurs in answer to a question, or a request, or a comment: *(Why don't you go?) Far. (Please call me when you go.) Fine. (I have a lot of good ideas but I don't know what to say.) Like that I also (=I feel just the same way)*. The Sequence type consists of a non-exclamatory predicate unaccompanied by a concomitant expression such as must occur in independent sentences.[6] It is a characteristic of Chontal narrative that such dependent predicates are found imbedded between normal independent sentences (which may be quite long and involved) as a type of summation or anticlimax. *(She was terrified. So white was her face that it had not a drop of blood.) Afraid. (The next day he was already ready where she would pass by.) Hid-himself.*

Not described in the Chontal monograph, but found in text are sequential sentences marked by sequence markers of both Chontal and Spanish origin. (In these examples, only the sequence marker is italicized.) *So* the man woke up. *But* every day he goes this way. *Therefore* they name their money. *And-so* he is very sick. *Thus* they call the money so that it will come.

In Isthmus Zapotec of Mexico, sequential, referential, and completive dependent sentences occur (Pickett, 1960, pp. 80–83). Referential sentences are not analysed as dependent in Pickett's framework, however, since they are not considered by her to be structurally different from independent types.

---

[6] Chontal non-exclamatory sentences consist of a noun or verb predicate plus some item to accompany the predicate, which I call a concomitant. The concomitant may be subject, but does not have to be. For discussion, see Waterhouse (1962, pp. 14–15). It is possible that further analysis of paragraph structure will make possible or necessary the analysis of sentences without subject or object as referential types as mentioned above.

Two types of sequential sentences are presented by Pickett: Class of sequence types, and Series. The first type is marked in one of three ways: (a) by a sequence introducer: *So then* the rabbit climbed on his back. *But* Saint Anthony was locked up. (b) By a double function sequence introducer: One day, *this fellow* was far from his home town. (c) By both: *And then, this chile* you wash. The second type is marked by the final dependent marker lá: *Then you dip it in egg* lá.

Completive sentences in Isthmus Zapotec include non-clause types and dependent clause types. Non-clause types are responses. These may be (a) declarative words or phrases: *(Are you going to market?) A little later.* (b) Yes-or-no interrogatives, marked by interrogative marker lá: *(Tomorrow bring that book.) This?* (c) Content-information interrogatives, marked by an interrogative word: *Who? How much?* Dependent clause types are either responses: *(When are you going to market?) As soon as I finish eating*; or parenthetical explanations: *Since he was a very rich man.*

For Tzeltal of Mexico eleven types of dependent sentences are described (Slocum, 1963). Two of these are functionally dependent: Direct Quotation and Response. (Except for the particles *yes* and *no*, responses in Tzeltal are full sentences.) The other nine types are all sequential, distinguished on the basis of the sequence marker which introduces them. Resultative is introduced by *therefore,* Sequential Action by *so*, Temporal Sequence by *afterwards*, Contrastive Subject by *on the other hand*, Contrastive Action by *however*, Conjunctive by *and*, Disjunctive by *but*, Explanatory by *truly* or *because*, Transitional by *well.*

From preliminary investigation of Witoto of Peru, it appears that omission of object is characteristic only of dependent transitive sentences; object is obligatory in independent transitive sentences.[7]

In Amuesha of Peru, three major types of dependent sentences are found (Duff, 1959, pp. 172–237): Introductory sentences, which begin the paragraph and anticipate something to follow, and include both Query and Introduction to Sequence types; Sequence sentences, which are found within the paragraph and are dependent upon some previous sentence in the same paragraph; and Response sentences, which include both responses to linguistic situations, such as answers to questions, and responses to non-linguistic situations, such as exclamations and greeting types. Both sequential and completive dependent sentences occur in the Amuesha materials.

Conclusions

Various ramifications of the concept of dependent sentences are found in various languages. In Zoque of Mexico (Engel and Longacre, 1963) and

---

[7] Data from Bryan Burton, Summer Institute of Linguistics.

Iquito of Peru (Eastman, 1963), structures of the sort I have called completive sentences are analysed as 'after-thoughts' which are considered part of the previous sentence, even though both the previous sentence and the afterthought are terminated with final juncture or sentence-final intonation. The choice of analysis in one or the other of these ways presumably rests on the importance of intonation contours to the analysis of sentences in the particular language.

Obviously, in languages like Greek, or Aztec,[8] or Zoque, where most sentences in a discourse after the first one are sequentially marked, there is little point in seeking to distinguish between independent sentences and dependent sequential sentences, unless the omission of sequence marker could be seen to mark the onset of a new unit, such as a new paragraph. The choice of sequence marker for a given sentence, however, might offer some clue to structure of higher-level units such as paragraphs, utterances, or episodes.

The concept of dependency on the sentence level is offered merely as a possibly useful tool for handling knotty problems of aberrant structures, and as a possible means for reaching analyses of levels higher than the sentence – a crucial problem in a number of aboriginal languages, especially when intelligibility demands translating connected texts as wholes rather than as strings of isolated sentences. It may also be useful in revising judgements of grammaticalness.[9] When more data are available, it may prove possible in many more languages to complete the morphosyntactic chain of dependencies all the way from morpheme to full discourse.

---

[8] Data from Earl Brockway, Summer Institute of Linguistics.

[9] See Maclay and Sleator (1960), in which they call their Type D meaningful and ordinary but ungrammatical; in the framework presented here, this type could now be called grammatical.

References

BLOOMFIELD, L. (1933), *Language*, Allen & Unwin.

DUFF, M. (1959), 'Amuesha (Arawak) Syntax I: Simple sentence types', *Publicações do Museo Nacional, Série Lingüística Especial*, no. 1.

EASTMAN, B. (1963), *Major Sentence Types and Clause Types*, unpublished manuscript.

ENGEL, R., and LONGACRE, R. E. (1963), 'Syntactic Matrices in Ostúacan Zoque', *Int. J. Amer. Linguistics*, vol. 29, pp. 331–44.

FRIES, C. C. (1952), *The Structure of English*, Longman.

KINDBERG, L. D. (1961), 'Independent and dependent sentence types in Campa (Arawak)', *A William Cameron Townsend en el Vigésimoquinto aniversario del Instituto Lingüístico de Verano*, Mexico City, Summer Institute of Linguistics.

LORIOT, J., and HOLLENBACH, B. (1970), 'Shipibo paragraph structure' *Foundations of Language*, vol. 6, pp. 43–66.

MACLAY, H., and SLEATOR, M. D. (1960), 'Responses to language: grammaticalness', *Int. J. Amer. Linguistics*, vol. 26, pp. 275–82.

NIDA, E. A. (1960), *Synopsis of English Syntax*, Norman.

PICKETT, V. B. (1960), 'The grammatical hierarchy of Isthmus Zapotec', Language dissertation no. 56, supplement to *Language*, vol. 36.

SLOCUM, M. C. (1963), *Tzeltal (Mayan) Sentence Structure*, unpublished manuscript.

WATERHOUSE, V. (1962), 'The grammatical structure of Oaxaca Chontal', publication no. 19, Indiana University Research Center in Anthropology, Folklore and Linguistics, *Int. J. Amer. Linguistics*, vol. 28, pt 2.

5 W. S. Allen

Transitivity and Possession[1]

W. S. Allen, 'Transitivity and possession', *Language*, vol. 40, 1964, pp. 337–43.

It is well known that in a number of European languages the perfect form of the verb employs an auxiliary identical with the verb used for predicating possession – e.g. English *I have loved*, French *j'ai aimé*, Portuguese *eu tenho amado*, Modern Greek *ékhō agapḗsei*. From discussions by Vendryes (1937) and by Benveniste (1952) it has also become clear that in transitive sentences the similarity between perfect and 'possessive' expressions is not confined to cases involving an auxiliary 'to have'.

In classical Armenian, for example, the transitive perfect is expressed by a participial form of the verb, with an auxiliary 'to be', and with the subject in the genitive case.[2] Thus, with an intransitive verb, *sa* (nom.) *ekeal ē* 'he has come'; but with a transitive verb *nora* (gen.) *teseal ē* 'he has seen'. The link with the 'possessive' here lies in the fact that the genitive is also used for predicating possession; thus *nora teseal ē* is formally parallel to *nora tun ē* 'he has a house'.[3]

---

[1] My thanks are due to Hans Vogt and John Lyons for their helpful comments on a preliminary draft of this paper.

[2] Discussions of this type often distinguish between a grammatical and a semantic subject. The former is formally definable and usually refers to the noun or pronoun which is in a particular (e.g. 'nominative') case and/or in concord with the verb or with a particular verbal element. The latter (also variously described as 'real' or 'psychological') generally remains undefined and intuitive, since situational or logical correlates such as 'actor' or 'topic' break down in a number of instances.

This double use of the term can lead to confusion. And since a grammatical category set up for one language cannot be identified as such with a grammatical category in another, the grammatical criterion is not valid for a cross-linguistic comparison. For such purposes it is the semantic criterion that is relevant, and 'subject' in the present study will therefore always imply semantic subject.

This admittedly poses theoretical problems of definition. But in the absence of a satisfactory solution there exists a formal correlate, such that the relative order of subject and object exponents in any one of the languages is preserved in translation into any other; and as Greenberg's investigations have shown (1963, p. 61), there is a virtual universal whereby (in sentences of neutral style and emphasis) a noun, pronoun, or noun phrase corresponding to the intuited semantic subject precedes any noun etc. corresponding to the semantic object (see also Jakobson in Greenberg, 1963, p. 212). By invoking this factor of order, we are of course introducing a form of grammatical criterion to fill out inadequacies of semantic definition – but one that has been shown empirically to have cross-linguistic relevance.

A similar situation is found in Old Persian, in sentences such as *ima tya manā krtam* 'this is what I have done'. As Benveniste observes, such constructions have commonly been treated as if they were passive transformations of an active sentence; and it is true that *krtam* is formally a passive participle. But *manā* is the dative-genitive case, whereas the agent in a passive transformation is expressed by the ablative. The genitive-dative case, on the other hand, is the case employed to express possession – e.g. *manā pitā* 'my father', or *kanbūjiyahyā brātā āha* 'Cambyses had a brother'. Thus in Old Persian also there is a connection between transitive perfect and 'possessive' expressions.

Something similar to the Old Persian is in fact found in ancient Greek, in examples such as (Hdt.) *hṓs moi próteron dedḗlōtai*, or (Dem.) *tà toútōi pepragména*, where the agent is expressed by the dative and not, as in a normal passive transformation, by the preposition *hupó* (Schwyzer and Debrunner, 1950, p. 150). This use of the dative is also paralleled in Latin by the old formula preserved in Varro, *sunt mi emptae*.[4] Thus, since in both languages the dative is used to predicate possession, a parallel exists between transitive perfect and possessive expressions even before the development of an auxiliary 'to have'.

Outside Indo-European one may find more general links between possessive and transitive, not confined to the perfect only. It is well known that in Hungarian, for instance, there is identity between some of the personal suffixes of the verb and possessive suffixes of the noun – e.g. in the definite form of the verb: *adom, adod, adja* 'I, you, he give(s)', beside *kalapom, kalapod, kalapja* 'my, your, his hat'.

Similar correspondences seem to be not uncommon in the American Indian languages. In Delaware, for example, 'A given form used in the expression of person serves to mark the possessor in the case of nouns, the actor in the case of verbs' (Voegelin, 1946, p. 143). In Yurok there is identity between the prefixes indicating possession and those indicating the subject of a verb (in subordinate clauses) (Robins, 1958, pp. 25, 50). Of the pronominal prefixes in Taos, Trager comments: 'In nouns the prefixes refer to the possessor and vary with the noun. . . . In verbs the prefixes refer to the subject and vary with the object' (1946, pp. 202, 209); and in a number of cases the nominal and verbal prefixes are formally identical (see Trager, 1954, pp. 173 ff.). In Kechua (Ayacucho dialect of Peru), Wonderly finds it appropriate to set up a single 'semantic component' to cover the 'effective mode (non-future or effected action) in verbs; possessional in nouns'; and here too one finds formal identity in some cases (1952, pp. 368 ff., 376). Suggestive also is Newman's comment on the Yawelmani dialect of Yokuts, that 'Ownership or possession is expressed by the possessive case suffix. . . . Agency in a passive construction is also indicated by the possessive case' (1946, p. 246). In some of these languages, however, the correspondence is not confined to transitive verbs.

---

[3] For further discussion see Benveniste (1959, pp. 57 ff.; 1960, p. 127).

[4] *RR* ii.2.5 'antiqua fere formula utuntur, cum emptor dixit: *Tanti sunt mi emptae* (sc. *oues*)? et ille respondit: *Sunt*'.

A particularly clear example of the transitive-possessive correspondence is presented by Eskimo, involving agreement both of case-form and pronominal suffixes. There is a system of two grammatical (as opposed to local) cases, 'absolute' and 'relative' (Swadesh, 1946, pp. 30 ff.; Schulz-Lorentzen, 1945) (also termed 'rectus' and 'versus' (Kleinschmidt, 1851), or 'intransitive' and 'transitive' (Bourquin, 1891)). To express possession, the noun indicating the possessor is put in the relative case, and that indicating the possessed takes a suffix agreeing with the possessor-noun in person and number; thus, *tiγiania-p iγlu-a* 'the fox's house' (lit. 'fox-rel. house-his'). To express transitivity, the subject is also put in the relative case, and the verb takes suffixes (of largely identical form) agreeing with the subject and object (Thalbitzer, 1930, pp. 1032, 1057); thus, *tiγiania-p iγlu takuβ-a-a* 'the fox saw the house' (lit. 'fox-rel. house saw-he-it'). With an intransitive verb, however, the subject stands in the absolute case.

The use of the relative case in Eskimo is summarized as follows by Swadesh (1946; see also Bourquin, 1891, p. 14):

> The relative is used when the substantive is in apposition to a 'possessor' pronominal reference attached to another noun, or to the subject in a double pronominal reference of a verb.

Similarly Schulz-Lorentzen:

> According to its meaning the relation borne by one word to another word may be a double one, either a relation of *ownership*, when it is governed by a noun form, or the relation of a *subject* when it is governed by a verb form with an object (1945, p. 18; see also Kleinschmidt, 1851, pp. 14, 69).

A further remarkable identity of expression is found in Northwest Caucasian. In Abaza, for example, there is a system of prefixes referring to the subject of an intransitive verb or the direct object of a transitive verb, and a system of infixes referring to the subject of a transitive verb and to the indirect object if any.[5] There is also a system of possessive nominal prefixes – e.g. *pa* 'son', *s-pa* 'my son', *l-pa* 'her son'. The systems of verbal prefixes and infixes, though partially similar to one another, show some significant differences, as may be seen from the following table (see Allen, 1956, p. 153). (The abbreviation 'irr.' means 'irrational'.)

---

[5] The relevant details apply equally to Abkhaz; see Deeters (1931, 289 ff.).

	Prefix	*Infix*		*Prefix*	*Infix*
sg. 1	*s-*	*-s-*	pl. 1	*h-*	*-h-*
2 masc.	*w-*	*-w-*	2	*š°-*	*-š°-*
2 fem.	*b-*	*-b-*	3	*y-*	*-r/d-*†
3 masc.	*d-*	*-y-*	relative	*y-*	*-z-*
3 fem.	*d-*	*-l-*			
3 irr.	*y-*	*-(n)a-**			

* *a* if it is the second element, *na* if it is later in the verbal complex.

† *r* is the basic form, with *d* as a morphophonemic variant.

The use of these elements may be illustrated by *aph°əspa aqac'a dlš'əd* 'the woman killed the man'; in the verb-form *dlš'əd* the prefix *d-* refers to the object (*aqac'a*) and the infix *-l-* to the subject (*aph°əspa*).[6] Similarly in *ylərtd* 'they gave it to her', the *y-* refers to the direct object, the *-l-* to the indirect object, and the *-r-* to the subject. Now the important point is that the system of verbal infixes is identical with the system of possessive prefixes (Genko, 1955, pp. 97 ff.); a form *d-l-š'ə-d* ('him-she-kill-ed') is structurally comparable with *d-l-pa-b* ('he-her-son-is'). The similarity of expression for transitivity and possession could hardly be closer (Deeters, 1931, pp. 297 ff.).

Abaza and Abkhaz have no case system. In the related Kabardian, however, there is a system of two cases;[7] and of these the so-called 'ergative' is in fact used both for the subject of a transitive sentence (and for the indirect object) and to express possession (Jakovlev, 1948, pp. 48 ff.). In Ubykh, with a similar system, the 'oblique' case, which indicates the subject of a transitive sentence (and the indirect object), is also used, though in the plural only, to express possession (Dumézil, 1959, p. 14; Vogt, 1963, p. 151).

The significance of such agreements is of course all the greater if the language has a larger number of cases. In the Northeast Caucasian language Lakk, for instance, there is a system of ten 'grammatical' cases[8] – and of these it is precisely the 'genitive' (or 'possessive') case that is used to express the subject of a transitive sentence (see Žirkov, 1955).

---

[6] It is true that in this as in many examples the 'infix' might be regarded as a second prefix. But since members of this system are frequently embedded between the elements of a compound verb (e.g. *y-š'tʼə-l-xd* 'she lifted it', from the verb *š'tʼe-x*), the term 'infix' is appropriate.

[7] Four cases are commonly attributed to Kabardian; but as Trubetzkoy notes (1939), two of these are really postpositional formations. See Jakovlev (1948, pp. 112 ff.).

[8] This number is agreed on (though with some difference as to their identity) by Dirr (1928) and Hjelmslev (1935). There are in addition a large number of 'local cases', of which the counts vary; Hjelmslev sets the total number of cases at forty-eight.

When we turn to South Caucasian, we find that such peculiarities are, as in the Indo-European languages, mainly confined to the perfect forms of the verb. The situation in Georgian is as follows. If the verb is in the present, the subject stands in the nominative case and the direct object in the dative; if the verb is in the aorist, the subject stands in the ergative (or 'narrative') case and the direct object in the nominative; and if the verb is in the perfect, the subject stands in the dative case and the direct object in the nominative. Thus,

kaci (nom.) *klavs megobars* (dat.)	'the man kills his friend'
kacma (erg.) *mokla megobari* (nom.)	'the man killed his friend'
kacsa (dat.) *mouklavs megobari* (nom.)	'the man has killed his friend'

The aorist construction has traditionally been interpreted as passive (e.g. by Schuchardt); but, as Vogt has shown, this is not a necessary interpretation (1950; see also Deeters, 1963). The perfect construction, however, is more generally agreed to be a passive transformation. Thus Vogt: 'Au parfait la conception du sujet et du régime grammatical est *invertie* . . . en d'autres mots, le parfait est un temps passif' (1938, p. 116); similarly Sommerfelt (1937); and Tschenkéli: '. . . werden die transitiven Verben in der 3. Zeitgruppe gewissermassen *passivisch* dargestellt' (1958).

Now it is true that the actual verb in the transitive perfect is passive in form, and that there is an inversion of the person markers – e.g. present *vc̣er* 'I write it' (where *v-* indicates first-person subject, and the third-person object is indicated by zero), perfect *(da)mic̣eria* 'I have written it' (where *m-* would elsewhere indicate a first-person (indirect) OBJECT, and *-a* a third-person SUBJECT). But it should be noted that the sentence is not entirely a passive transformation; for the agent in such a transformation is normally expressed by the instrumental case or by a postposition (*-mier*), as in *saxli šendeba mamis-mier* 'the house is built by the father' (beside active *mama ašenebs saxls*).[9] The subject of the transitive perfect, on the other hand, stands in the dative case; and the word order remains unaltered from the active.

If one takes the typical form of a transitive perfect without preverb, it is in fact identical not with a 'passive of process' but with a so-called 'passive of state', in the 'objective version'.[10] Thus a form such as *mic̣eria* may be either the perfect of a transitive active verb or the present of a 'passive of state', and so is translatable according to context (Tschenkéli, 1958, p. 509) either as 'ich habe etwas geschrieben' or 'etwas ist/steht bei mir geschrieben'. And what is of particular relevance is the semantic relationship between the objective version of the passive of state and its indirect object; of the person markers in such verbs Tschenkéli states (1958, p. 437): 'sie geben im allgemeinen nur einen Hinweis auf ein indirektes Objekt oder *deuten ein Besitzverhältnis an.*' Thus, as Vogt points out (1938, p. 149), a sentence like *cxeni mibia*, lit. 'the horse is tied up for me' (where *mibia* is the passive of state of the verb *b*, and *m-* indicates the first-person singular indirect object) may be paraphrased as *cxeni mqavs mibmuli* 'I have the horse tied up' (where *mqavs* is 'I have', and *mibmuli* is the past participle passive). Since the person markers of the indirect object in the passive correspond to those of the subject in the transitive perfect active, a connection is thus established between the latter and expressions of a possessive nature.[11]

---

[9] Vogt (1938): 'l'agent au passif s'exprime par l'instrumental s'il est un nom de chose, par postposition si c'est une personne'.

[10] The 'versions' of the Georgian verb express particular nuances of semantic relationship between subject and object(s).

Note incidentally that in Georgian the dative case and 'indirect object' person markers are also used for the subject of a number of common verbs in the present, notably the so-called 'Empfindungsverben' (and including the verbs meaning 'to have') (see Chanidzé, 1963, pp. 1 ff.; and Deeters, 1963, p. 61).

There remains to mention in Georgian a peculiarity connected with the indirect object, which is of some interest both in itself and as having a remarkable parallel in another, unrelated language which will be discussed. When the verb is in the present or aorist, the indirect object is in the dative. In the present, this means that both direct and indirect object are in the dative, and for this reason Tschenkéli refers to it as the 'dative-accusative'. But when the verb is in the perfect, and the subject is accordingly in the dative, the indirect object is expressed by a postposition *-tʿvis*, which is normally translatable as 'for' (e.g. *es çerili aris dedisa-tʿvis* 'this letter is for mother'). Thus we have

PRESENT: *studenti* (nom.) *sçers megobars* (dat.) *çerils* (dat.) 'the student writes a letter to his friend';
AORIST: *studentma* (erg.) *misçera megobars* (dat.) *çerili* (nom.);
but:
PERFECT: *students* (dat.) *miuçeria megobrisa-tʿvis* (ppn.) *çerili* (nom.).

Thus, where the dative case is preempted by the subject, the indirect object must be expressed by other means.[12] This seems to argue a special relationship between the subject and the indirect object in such sentences.

---

[11] A more obvious connection, reminiscent of Old Persian, is seen in participial constructions of the type *čemi daçerilia es çigni*, lit. 'my written-is this book', i.e. 'this book has been written by me' (Vogt, 1938, pp. 254 ff.).

[12] This applies even where the subject is not actually expressed by a nominal or pronominal form, but is implied by the verbal person marker; e.g. *mimiçeria çerili mis-tʿvis* 'I have written him a letter'.

With transitive verbs the indirect object can hardly be said to indicate a state of possession; but it does in most cases indicate a process of 'receiving' or 'acquisition', i.e. of 'coming into possession'. (It may be recalled that in Abaza, Kabardian, and Ubykh there are identical modes of expression for possession and indirect object.) And the relationship between this process and 'perfect transitivity' is such that, although in different sentences they are identically expressed, within a given sentence they must have different expressions.

The final example is taken from Indo-Aryan, and concerns the Harauti dialect of Rajasthani.

A number of the modern Indo-Aryan languages have developed special modes of expression for the subject of transitive verbs in the perfect and aorist; these differ formally from one another only in the presence or absence of an auxiliary 'to be', and are sometimes referred to by the joint title of 'perfective'. Thus in Hindi the subject of an intransitive verb, or of a non-perfective transitive verb, stands in the direct case; e.g. *lərka dəwṛta/dəwṛa həy* 'the boy runs/has run'; *lərka roṭi khata həy* 'the boy eats bread'. But if the verb is transitive and perfective, the subject stands in the oblique case with a postposition *-ne*; e.g. *lərke-ne roṭi khai* 'the boy ate bread'. The indirect object is expressed by a postposition *-ko*; e.g. *məyŋ-ne lərke-ko roṭi di həy* 'I have given the boy bread'. In transitive sentences with a perfective verb, there is concord between the verb and the direct object, which stands in the direct case; and this fact has led many writers to interpret the construction as passive (Kellogg, 1938, paras. 412, 785). There is indeed historical support for this, but such an interpretation is not descriptively appropriate (Allen, 1950–51, pp. 71 ff.).

A similar situation is found in Gujarati; but here the transitive subject of a perfective verb has a special case-suffix *-e*; e.g. *chokre paṭh vaŋcyo* 'the boy read the lesson'; and the indirect object is expressed by a postposition *-ne*; e.g. *iśvəre teo-ne dikri api che* 'God has given them a daughter'.

Harauti stands geographically and linguistically between Hindi and Gujarati. And one aspect of this dual affinity consists in its having a postposition *-nɜ* both for the transitive subject of a perfective verb and for the indirect object. Thus *chora-nɜ photi phəḍi chɜ* 'the boy has read the book'; *muŋ chora-nɜ photi duŋgo* 'I will give the boy a book'. The two *nɜ*'s, however, like the Georgian datives, are mutually exclusive (Allen, 1960, p. 10). If the transitive verb is in the perfective, the subject takes the postposition *-nɜ*, and the indirect object must then be expressed by a periphrasis, by means of the double postposition *-kɜ-taiŋ*; e.g. *mhə-nɜ chora-kɜ-taiŋ photi di chɜ* 'I have given the book to the boy'. In other contexts *-kɜ-taiŋ*, like the Georgian *-tʻvis*, has the meaning 'for'; e.g. *choro chori-kɜ-taiŋ kam kərɜ chɜ* 'the boy does the work for the girl'.

Thus, precisely as in Georgian, there is a relation of mutual exclusiveness between the transitive subject and the indirect object; and the situation is resolved in precisely the same way. The comments made on this relationship in Georgian therefore apply equally to the Harauti situation.

The Harauti peculiarities in regard to *-nɜ* were discussed in a paper presented to the summer meeting of the Linguistic Society of America in 1953 (Allen, 1954, p. 4). In view of their mutually exclusive distribution, I then considered the possibility of treating this form in its different functions as instances of the same morpheme (as opposed to homonymous morphemes), but rejected this on the grounds of *a priori* semantic incompatibility. The findings of the present study, however, provide empirical support for such an interpretation.

In a variety of languages a connection is clearly established between the expression of the subject, more particularly the transitive subject, and of 'possession' or 'acquisition', the connection being in some cases limited to perfect or perfective forms of the verb. The modes of expression differ widely, but serve only to emphasize the underlying parallelism of the phenomena.

References

ALLEN, W. S. (1950–51), 'A study in the analysis of Hindi sentence-structure', *Acta Linguistica Holmiensia*, vol. 6, pp. 68–86.

ALLEN, W. S. (1954), 'A problem of morpheme identity in a Rajasthani dialect', *Linguistic Soc. Amer. Bull.*, vol. 27, no. 4.

ALLEN, W. S. (1956), 'Structure and system in the Abaza verbal complex', *TPhS*, pp. 127–76.

ALLEN, W. S. (1960), 'Notes on the Rajasthani verb', *Indian Linguistics*, vol. 21, pp. 4–13.

BENVENISTE, E. (1952), 'La construction passive du parfait transitif', *Bull. de la Soc. de Linguistique de Paris*, vol. 48, pp. 52–62.

BENVENISTE, E. (1959), 'Sur la phonétique et la syntaxe de l'arménien classique', *Bull. de la Soc. de Linguistique de Paris*, vol. 54, pp. 46–68.

BENVENISTE, E. (1960), 'Être et avoir dans leurs fonctions linguistiques', *Bull. de la Soc. de Linguistique de Paris*, vol. 55, pp. 113–34.

BOURQUIN, T. (1891), *Grammatik der Eskimo-Sprache . . . an der Labradorküste*, London, Moravian Mission Agency.

CHANIDZÉ, A. (1963), 'Le sujet grammatical de quelques verbs intransitifs en géorgien', *BSL*, vol. 58, pp. 1–27.

DEETERS, G. (1931), 'Der abchasische Sprachbau', *Nachr. von der Ges. der Wiss. zu Göttingen, Phil.-Hist. Kl.*, pp. 289–303.

DEETERS, G. (1963), 'Die kaukasischen Sprachen', *Handbuch der Orientalistik*, Abt. 1, vol. 7, 'Armenisch und Kaukasische Sprachen', pp. 1–79.

DIRR, A. (1928), *Einführung in das Studium der kaukasischen Sprachen*, Leipzig, Verlag der Asia Major.

DUMÉZIL, G. (1959), *Etudes oubykhs*, Paris, Maisonneuve.

GENKO, A. N. (1955), *Abazinskij jazyk*, Moscow, Izd-vo Akademii Nauk USSR.

GREENBERG, J. H. (1963), *Universals of Language* MIT Press.

Hjemslev, L. (1935), *La catégorie des cas*, Copenhagen, Aarhus University Press.

Jakovlev, N. G. (1948), *Grammatika literaturnogo kabardino-čerkesskogo jazyka*, Moscow, Izd-vo Akademii Nauk USSR.

Kellogg, S. H. (1938), *Grammar of the Hindi Language*, Trubner & Co.

Kleinschmidt, S. (1851), *Grammatik der grönländischen Sprache*, Berlin, G. Reimer.

Newman, S. S. (1946), 'The Yawelmani dialect of Yokuts' in L. Bloomfield, H. Hoijer, *et al* (eds.), *Linguistic Structures of Native America*, pp. 222–48.

Robins, R. H. (1958), *The Yurok Language*, University of California.

Schulz-Lorentzen, C. W. (1945), *A Grammar of the West Greenland Language*, Copenhagen, C. A. Reitzel.

Schwyzer, E., and Debrunner, A. (1950), *Griechische Grammatik*, vol. 2, *Syntax*, Münich, Beck.

Sommerfelt, A. (1937), 'Sur la notion du sujet en géorgien moderne', *Mélanges van Ginneken*, pp. 85–92.

Swadesh, M. (1946), 'South Greenlandic Eskimo', in L. Bloomfield, H. Hoijer *et al.* (eds.), *Linguistic Structures of Native America*, pp. 30–54.

Thalbitzer, W. (1930), 'Eskimo', *Handbook of American Indian Languages, Bureau of American Ethnology Bulletin*, vol. 40, pp. 967–1069.

Trager, G. (1946), 'An Outline of Taos Grammar', in L. Bloomfield, H. Hoijer *et al.* (eds.), *Linguistic Structures of Native America*, pp. 134–221.

Trager, G. (1946), 'Taos II: Pronominal reference', *Int. J. Amer. Linguistics*, vol. 20, pp. 173–80.

Trubetzkoy, N. (1939), 'Le rapport entre le determiné, le determinant et le defini', *Mélanges de linguistique offerts à Charles Bally*, pp. 75–82. Reprinted in R. Austerliz, E. Hamp and F. W. Householder, *Readings in Linguistics*, vol. 2, 1966, Chicago University Press, pp. 133–8.

Tschenkéli, K. (1958), *Einführung in die georgische Sprache*, Zürich Amirani Verlag.

Vendryes, J. (1937), 'Sur l'emploi de l'auxiliare "avoir" pour marquer le passé' *Mélanges van Ginneken*, pp. 85–92.

Voegelin, C. F. (1946), 'Delaware, an Eastern Algonquian Language', in L. Bloomfield, H. Hoijer, *et al.* (eds.), *Linguistic Structures of Native America*, pp. 130–57.

Vogt, H. (1938), *Esquisse d'une grammaire du géorgien moderne*, Oslo, A. W. Brøgger.

Vogt, H. (1950), 'Un aspect du problème actif-passif dans le verbe', *Journal de Psychologie*, vol. 43, pp. 130–38.

Vogt, H. (1963), *Dictionnaire de la langue oubykh*, Oslo, Universitetsforlaget.

Wonderly, W. L. (1952), 'Semantic components in Kechua person morphemes' *Language*, vol. 28, pp. 366–76.

Žirkov, L. I. (1955), *Lakskij jazyk: Fonetika i morfologija*, Moscow, Izd-vo Akademii Nauk USSR.

Part Two
Categories and Arrangements

The relation of a taxonomic emphasis to the problems of describing strange languages is clear from the fact that over half of these selections are concerned with particular languages (other than English) or language families. Several of them are exemplars of the methods of specific 'schools'. Reading 17 by Robins in Part Three or Reading 20 by Honey in Part Four might have been shifted here to represent Firthian practice. The influence of Chomsky appears here in almost all the articles after 1960, but in particular in Yngve.

The typical pre-Chomsky emphasis on language variety, the unique rather than the universal, is conspicuous in many of the selections.

6 L. Bloomfield

Outline of Ilocano Syntax

L. Bloomfield, 'Outline of Ilocano syntax', *Language*, vol. 18, 1942, pp. 193–200.

1. The syntactic EXPRESSIONS (words and phrases) of Ilocano are divided by their functions, first of all, into OPEN expressions and OBJECT expressions: open: *simrek* 'he entered'; object: *ti simrek* 'one who entered'.[1]

2. Words fall into three main classes:
PRONOUNS, which are always object expressions: *syak* 'I';
FULL WORDS, which by themselves are open, but can be objectified by certain proclitic particles: *simrek*;
PARTICLES, which are open and are never objectified: *ken* 'to, and', *ti*.

3. Object expressions are divided into two categories, STRAIGHT and OBLIQUE. The straight category is subdivided into NOMINATIVE and

---

[1] Our transcription uses the following symbols: *a e i u ʔ p t k b d g m n ŋ l r s*. In exclamations and foreign words there are also *h* and *c* (much as in English *church*); in foreign words also *j* (much as in English *judge*) *f v o*. Further, we use *y, w* for nonsyllabic positional variants of *i, u*.

The accent mark over a vowel indicates lengthening, high pitch, and stress. Words and phrases which contain no vowel of this kind have a weaker and unstable stress (without lengthening or marked rise of pitch) usually on the last or next to last vowel indifferently, but in some instances on an earlier vowel; in no case are they stressed on a vowel followed by single consonant plus vowel; thus, *napintas* 'beautiful' is weakly stressed on the last or next to last vowel, indifferently, and *silid* 'room', *daga* 'earth, land' are weakly stressed on the last vowel.

We mark open juncture by space or hyphen. The phrase finals (. ; :) and sentence finals (. ? !) are well marked, especially by features of pitch; before them a word or phrase with indifferent stress usually has its weak stress on the last vowel.

Where we write no consonant before an initial vowel or between vowels, a weak variety of the glottal stop (varying toward zero or *y, w*) is to be understood: *ásu* 'dog', *naimbag* 'good'; contrast *dáʔit* 'sewing', *salʔit* 'lightning', *agaʔʔárak* 'smelling of rum', with stable glottal stops.

Our data were obtained from Mr Severino Montano of Laoag, Ilocos Norte, a student in the Yale School of Drama. Apart from errors of recording, they suffer from the defects that are inevitable when speech forms are presented *ad hoc* and when texts are dictated.

Needless to say, this sketch gives only the main outlines of the syntax and omits many interesting features. It represents a part of studies in Malaysian (and related languages) carried on at Yale University under the auspices of the National School of Modern Oriental Languages and Civilizations.

ATTRIBUTIVE, but this subdivision appears only in the enclitic pronouns.

Object expressions are firstly, pronouns, and, secondly, full words and phrases OBJECTIFIED by certain proclitic particles.

4. ENCLITIC PRONOUNS appear in straight form only; they alone distinguish between nominative and attributive forms:

	NOMINATIVE	ATTRIBUTIVE
‘I’	*ak*	*ku*
‘thou’	*ka*	*mu*
‘thou and I’	*ta*	*ta*
‘we, inclusive plural’	*tayu*	*tayu*
‘we, exclusive’	*kami*	*mi*
‘ye’	*kayu*	*yu*
‘he, it’	(zero)	*na*
‘they’	*ida*	*da*

These forms are spoken in close juncture with a preceding word, except that in certain cases where two are present the second may be added with open juncture.

The enclitic pronouns are subject to various habits of sandhi. Thus, *ku* and *mu* after vowels (with certain exceptions) lose their *u*: *ti balayku* ‘my house’, but *tuy ímak* ‘my hand’. After *di* ‘not’ they do not lose their vowel, but *ku* is here usually replaced by *ak*: *dyak nakíta* ‘not by-me (it) has-been-seen; I have not seen it’, beside less usual *díku nakíta*. In most connections, but not for instance after an attributive enclitic, *ida* drops *i*: *inismanmi ida* ‘smiled-at by-us they; we smiled at them’, but *simrekda* ‘they entered’.

Among these and similar modifications, the most striking are those which occur when an attributive and a nominative enclitic (the order is always this) come together.

The form for ‘by-me thou’ omits *ku*: *inismanka* ‘smiled-at thou; I smiled at thee’;

ak loses its vowel after *na* and *da*: *inismannak* ‘he smiled at me’, *inismandak* ‘they smiled at me’.

When the first two persons come together, the attributive forms of the third person are used (except for ‘by-me thou’, above): *inismannak* ‘thou didst smile at me’; *inismandak* ‘ye smiled at me’; *inismanna kami* ‘he smiled at us’ and ‘thou didst smile at us’; *inismanda kami* ‘they smiled at us’ and ‘ye smiled at us’; *inismandaka* ‘they smiled at thee’ and ‘we smiled at thee’.

5. The non-enclitic pronouns distinguish between straight and oblique forms. Those of the first and second persons are derived from the nomina-

tive enclitic forms by means of proclitic elements *si, ni, da, ken, ka*. The third person straight forms are based on a word *isu* 'he, it', which in certain respects resembles a full word, making, for instance, the reduplicated *is ʔisu* 'only he'. The oblique third person forms are phrases containing the full word *kwa* 'property'.

	STRAIGHT	OBLIQUE
'I'	*syak*	*kanyka*
'thou'	*sika*	*kenka*
'thou and I'	*data*	*kadata*
'we, inclusive plural'	*datayu*	*kadatayu*
'we, exclusive'	*dakami*	*kadakami*
'ye'	*dakayu*	*kadakayu*
'he, it'	*isu*	*ken kwána*
'they'	*isuda*	*kada kwáda*

6. We come now to the proclitic particles which objectify full expressions.

Among full words, a special sub-class of PERSONAL NAMES is objectified by the following particles:

straight *ni*: *ni pedru* 'Pedro', *ni ama* 'Father';

oblique *ken ni*: *ken ni pedru, ken ni ama*;

straight plural *da*: *da patrya ken ni kamílu* 'Patria and Camilo', *da baket pemya* 'old Femia and her people';

oblique plural *kada*: *kada patrya ken ni kamílu*.

More usually the plural is replaced by expressions like *ni hwan ken ni pedru* 'Juan and Pedro'.[2]

7. Full forms which are not names are objectified for the straight form by the following six proclitic particles:

ti indifferent, including mere anaphora: *ti libru* 'the book, a book';

tuy 'this, near the speaker': *tuy libru* 'this book here';

ta 'that', especially when near the addressee: *ta libru* 'that book, that book you have there';

dyay 'yon' decidedly far off, also poetic for *ta*: *dyay lakay* 'that old man';

tay 'that', not in sight, but known and extant: *nalipátanna tay librúna* 'He has forgotten that book of his';

di 'that', no longer extant: *di dáʔan ŋa simbáʔan* 'the old church (that used to be here)'.

The oblique forms corresponding to these are made by prefixing *i-*: *iti, ituy, ita, idyay, itay, idi*. These forms, however, are subject to various habits of sandhi, including loss of *i-*: *nakaálaʔak ti ikan* 'I have caught some fish, a fish', beside less usual *nakaálaʔak iti ikan.*

---

[2] Mr Montano remarked that these plurals with *da* are never used in the presence of the persons named.

Non-proclitic uses of some of these forms appear in paras. 20, 31, 32, 34 (see pp. 98–102); in general, however, only the forms in paras. 8, 9 appear as independent object expressions.

8. A set of more emphatic forms is made by prefixing *dey-*: *deyti*, *deytuy*, *deyta*, *deydyay*, *deytay*, *deydi*. The form *deydyay* is common, in contrast with *dyay* above: *tawágem deydyay ubiŋ* 'Call that child over there'. These longer forms are also used as (non-proclitic) object expressions: *deytuy* 'this (person or thing) here'.

The oblique of these emphatic forms is made by prefixing *iti*: *pukánem deyta káyu iti deytuy ŋa buneŋ* 'Cut that tree with this bolo'. However, there are also two forms with *di-*, specialized in local meaning: *dituy* 'in this, here' and *dita* 'in that, there'; for the farther demonstrative the form *idyay* is in parallel use.

9. A set of plurals is made by prefixing *dagi-*: *dagiti*, *dagituy*, *dagita*, *dagidyay*, *dagitay*, *dagidi*. These are used both as proclitic objectifiers and as object expressions: *dagituy tattáʔu* 'these people', *dagituy* 'these (persons or things)'. Like all plurals (except those of the personal pronouns), these are not categoric: singular forms are freely used where more than one object is involved: *ti ikan* may mean one or more than one fish; plurals of varying explicitness are *dagiti ikan*, *ti ikʔikan*, *dagiti ikʔikan*; the forms with pluralized (here reduplicated) full word are used especially where more than one set or variety is meant.

The plural objectifiers make oblique forms with proclitic *ka*: *kadagiti*, and so on.

10. The most usual construction of full sentences consists of an open expression containing a full word, followed by a straight object expression: *mabisin deyta ubiŋ* 'That child is hungry'. We shall describe these two positions as PREDICATE and SUBJECT. Pronouns appear here in nominative enclitic form: *mabisínak* 'I am hungry', *áriʔak* 'I am a king'. Since the third-person singular in this category is zero, any open expression in principle makes a full sentence: *mabisin* 'He is hungry'.

11. Certain open expressions, centering round certain full words which we shall call PASSIVES, are accompanied by two straight object expressions: *nakíta ni marya ti ári* 'Was-seen by-Maria the king; Maria saw the king'. Enclitic pronouns are attributive in the former position, nominative in the latter: *inismánna ida* 'Smiled-at by-him they; He smiled at them', *nakítak* 'I saw him' (*k* for *ku*, zero for third-person singular nominative). We call the prior (attributive) straight object expression the AGENT in this construction.

12. The subject and also, with passives, the agent, is often anticipated by an object expression with phrase final rise of pitch (marked by comma) before the predicate: *dagituy ubbiŋ*, *mawwawda* 'The children, they are thirsty'. This is especially common where the subject is third-person singular: *deytuy ubiŋ*, *mabisin* 'This child, it is hungry'; *isu*, *naglwálu* 'He, he prayed'; *ni ugsa*, *nalipátanna ti lumba* 'The stag, he (agent) had forgotten about the race'. Sometimes an enclitic nominative is then left off: *syak*, *áriʔak* 'I, I am a king', but also *syak*, *ári*.

13. Full sentence constructions consisting of open expressions are decidedly limited; the following deserve mention.

A predicate-subject construction is that of *adda* 'there is' followed by an open expression: *adda bagas* 'There is rice'. Contrast *adut bagas* 'There is lots of rice', where *t* is a sandhi form for *ti*; also *awan ti bagas* 'There is no rice'.

The proclitic word *inn* 'go' (e.g. *inka* 'Go thou') is combined with certain open forms; it then means 'where?' as *impapanam?* 'Where are you going?' with *papanan* 'place gone to' and *m* as sandhi form for *mu*.

Certain words with attributive pronouns or other attributes form full sentences: *naggapwam?* 'Where do you come from? Where are you coming from?' with *m* for *mu*; *nagpúduten* 'How hot it is!'

14. A second important type of full sentence construction consists of two object expressions with phrase-final after the prior one: *deytuy*, *ti silid ŋa pagadʔadálak* 'This is the room where I always study'; *syak*, *ti maŋ patay* 'I am the one who can kill him'. Often *isu* is prefixed to a third person after the phrase final: *ti pagsúratmu*, *isu deytuy plúma* 'What you are to write with is this pen'.

15. The chief types of attributes are: straight object expressions (paras. 16–19, see pp. 97–8), oblique object expressions (paras. 20–22, see p. 98), open expressions with proclitic *ŋa* (paras. 23–5, see pp. 98–9), enclitics (para. 29, see p. 100), and phrasal attributes introduced by various particles and full words (paras. 30–34, see pp. 100–101).

16. Straight object expressions serve as attributes of object expressions and denote a possessor: *ti ípus ti nwaŋ* 'the tail of the carabao'. Enclitic pronouns are in attributive form: *ti balayna* 'his house', *tuy ímak* 'my hand'.

These attributes are sometimes anticipated by a straight object expression: *ni nwaŋ*, *nakakipkípet ti báduna* 'The carabao, very tight the clothes of-him; The carabao's clothes are very tight'.

17. Other straight attributes occur only in the shape of certain special phrases: *agádalkat saŋkabassit* 'Study thou a little while'; here *t* is a sandhi form for *ti*, and *ti saŋkabassit* 'a little while' is a special attributive expression; similarly *ti asideg* 'near by', *ti masápa* 'early', and some others.

18. The passive *ikkan* 'given to', with its derivatives, takes not only straight forms for agent and subject, but optionally a third straight form for that which is given: *ikkannak met ti* (or *iti*) *saba* 'Give me too some bananas'.

19. Certain open expressions, most of them passive, are modified by a straight object expression which names a hypothetical state or further action of the agent: *kayatkut aglaŋuy* 'I want to take a swim'; here *t* is a sandhi form for *ti*, which objectifies *aglaŋuy* '(he) swims' (*aglaŋúyak* 'I swim'); *kayatku* '(It is) desired by-me' is a passive. Similarly *nalipátanmi ti napan idyay eskwéla* 'Forgotten by-us (objectifier) went to school; We forgot to go to school'. Compare with these attributes the subject of a sentence such as *nasakit ti agparparintúmeŋ ŋa mabayag* 'Pained (objectifier) kneeling a long-time; It is painful to be kneeling down a long time'.

20. Oblique object expressions as attributes are very common and have many meanings, such as place, instrument, and so on: any object expression that is not subject, agent, or possessor, is oblique: *pukánem deyta káyu iti deytuy ŋa buneŋ* 'To-be-cut by-thee (Cut thou) this tree with this bolo'; *iti ámin ŋa babbalásaŋ, ni kunswélu, ti kapintásan* 'Of all the young women, Consuelo is the prettiest'; also *ni kunswélu, ti kapintásan kadagiti ámin ŋa babbalásaŋ*. Other examples are *ita ŋa aldaw* 'today', *idi kalman* 'yesterday', *itay bigat* 'this morning (past)'.

So, especially, the forms *itan* 'now', *idin* 'then': *ó, naálaʔak itan* 'Oh, you've caught me this time'. Here *ita* and *idi* appear in non-proclitic use, against para. 7 (p. 95); for the *n* see para 29 (p. 100).

21. To passives there generally correspond non-passive forms whose subject corresponds to the agent of the passive; with these, an oblique object expression, denoting usually an indefinite thing, corresponds to the subject of the passive: *nakakita iti meysa ŋa bassit ŋa balay* 'He saw a little house': the corresponding passive would be *nakítana ti balay* 'He saw the house'. We give the name GOAL to the oblique expression which corresponds to the subject of a passive.

22. A few objectifiers have a special form for the local oblique (para. 8, p. 96): *aglaŋlaŋuykamit dituy púlug* 'We swim in this lake'; here *kamit* is a sandhi-form for *kami*.

23. Attributes with the proclitic particle *ŋa* may be called CONJUNCT; in general they imply identity of attribute and head: *ti meysa ŋa napintas ŋa tumaytayab* 'a pretty bird', where *meysa* 'one' and *napintas* 'pretty' are attributes. Some of the objectifiers, especially *deytuy* and *dagituy* are often treated as conjunct attributes instead of proclitics: *deytuy* (*ŋa*) *buneŋ* 'this bolo'.

24. However, many other kinds of attributes are joined with *ŋa*; *agparparintúmeŋ ŋa mabayag* 'He kneels for a long time'. So especially *saʔaŋ* 'not',

which precedes its head as a CLOSE attribute; by this we mean that it is followed by the enclitics which would otherwise follow the head: *sa ʔanmu ŋa liplipátan deyta kalluguŋmu* 'Not by-thee to-be-forgotten that hat of-thee; Don't forget your hat'.

25. Similarly, the passive *kayat* comes as a close conjunct attribute before passives, where it may be said to modify the subject: *kayatka ŋa aŋgwen* 'Desired (by-me) thou as one-who-is-kissed; I want to kiss you'; compare *sika, ti kayatky ŋa kaay ʔáyam* 'Thou art the desired by-me person-played-with; You are the one I want to play with'.

Somewhat differently, a non-passive open expression for a hypothetical state of the subject of *kayat* follows it as a conjunct attribute: *kayatnak ŋa aggubernadur* 'Desired by-him I as one-who-is-governor; He wanted me to be a governor'. This supplements the construction of para. 19, p. 98.

26. Unmarked open expressions as attributes are less numerous. An important one is *di* 'not', which is a close attribute: *díka umasʔasideg ita danum* 'Do not go near the water'; *dímu basbasaʔen deyta bádum* 'Do not get thy clothing wet'.

Similarly *sa* 'then': *sáda naglaŋuy* 'Then they went swimming', where the enclitic subject *da* for *ida* is drawn away from *naglaŋuyda* 'They went swimming'.

27. Certain words serve as unmarked attributes; so especially *únay* 'much, many', which follows its head: *dagiti bradlis, nasúdi únay ŋa gagayyemmi ida* 'The Bradleys, they are very dear friends of ours'. Similarly *láʔeŋ* 'only': *bayʔam láʔeŋ* 'Just leave him be; Never mind him'; in an object expression it precedes: *ti láʔeŋ pagsásaʔúda* 'their only way of communicating'. Also *ŋarud* 'so, therefore, accordingly': *ŋarud, ni manwel lópes, imbílaŋna ŋa kabsat ni mister bradli* 'So then Manuel Lopez looked upon Mr Bradley as a brother'; *maŋrugí ŋarud ti stórya* 'So now the story begins'; in this example *ŋarud* can be described as an enclitic (see para. 29, p. 100).[3]

Other words are sometimes unmarked and sometimes conjunct: *patyek ámin ti ibagbagáda* 'I believe everything they say'; *isu ti agraraman ámin kada kwáda* 'He is the most demanding of them all', but *ámin* conjunct in an object expression in para. 20, p. 98. Similarly, *kankanáyunna ibábaʔun ni deyvid idyay pagadálan* 'She is always sending David to school', where *kankanáyun* 'always' is a close attribute, attracting the agent *na* from the passive *ibábaʔun* 'is being sent'; but also *kankanáyunna ŋa malpas ti masápa* 'He always finishes it early' and, not close, *kankanáyun ŋa malipátak deytuy libru* 'I always forget this book', where *ku* is not drawn away from the passive *malipátan*.

---

[3] In this and some of the following examples, some final vowels (*maŋrugí*, *intú*, *dí*, *lá*) were recorded with the high stress and lengthening which are peculiar to open syllables: this implies close juncture with the following word, and properly our transcription should leave no space. However, we have left spaces before all words which seem to be preceded prevalently by open juncture.

Other unmarked attributes are phrases: *di mabayag* 'not long afterwards; then soon', in contrast with *mabayag* 'for a long time', which is conjunct: *di mabayag, ginúyud ni paduŋ ti ípus ni suŋgu* 'After a little while, Mr Turtle pulled Mr Monkey's tail'.

28. An important type of unmarked attribute consists of certain words which denote a performer of locomotion; they are close attributes: *mapantu agádal* 'He goes (enclitic of future time) he studies; He will go to school'; *tay ulitegku, immayna kami sinaruŋkáran* 'That uncle of-me, (he) came by-him we were-visited; My uncle came to visit us'; here the enclitics *na kami* are drawn away from the passive *sinaruŋkáranna kami* 'He visited us'.

29. There is a considerable variety of enclitic attributes. Some of them are added always in close juncture; others seem to vary. Thus, *tu* marks future time: *agádalaktu* 'I shall study'; a preceding vowel adds *n*: *agádaldantu* 'They will study', but *tayu*, *kayu* drop *u*: *agádalkaytu* 'Ye will study'. *en* 'already, now', as in *nagáwiden* 'He has gone home', drops its vowel after a vowel: *natúrugdan* 'They have already gone to bed', and *ak* adds *u* before it: *matáyakun* 'I am already as good as dead'; *ku* and *mu* after a vowel keep their *u* before *en*: *irugímun* 'Begin now', beside *irugim* 'begin'; where *irugi* '(to be) begun' is passive, with agent *mu*. Others are *la* 'only', *man* 'do, please' (*manen* 'also'), *met* 'also', *pay* 'still', *ʔa* 'don't you agree? see here!', *ʔaya* 'isn't it so?', *kuma* 'hypothetical, ought', *kanu* 'quotative'.

There is no hard and fast line between these enclitics and the absolute non-enclitic attributes of para. 26, p. 99: we class as non-enclitic those which at times come first in a phrase or sentence, such as *ŋarud* in para. 26, p. 99.

30. Phrasal attributes are headed by certain words, mostly proclitic particles, and are added without connectives. One could say that attributes with *ŋa* and object expressions with *ti* and so on belong in this type.

The resemblance of *ken* 'and' to coordinators of other languages does not reach far: *ti gubat dagiti filipínus ken dagiti kastílas* 'the war of the Filipinos and the Spaniards'; but compare *ken* in paras. 5, 6, p. 95: *nagsúratak ken ni inak* 'I have written a letter to my mother'. Predicative expressions are headed by *ket* 'and' and *ŋem* 'but': *immasidégak, ket tinalyawnak, ŋem timmayab* 'I went near and it looked at me, but it flew away'; compare, however, *kartíŋam, ket marúnaw ti pagpalamiʔis* 'Hurry up, or the ice cream will melt', and *ŋem* 'than': *nabanbannúgak ŋem isu* 'I am more tired than he'.

31. Comparable to straight object attributes are clauses with *ta* 'because, that': *immáyak dituy, ta kayatku ŋa kasaʔu ni senyóra bradli* 'I have come here because I want to speak to Mrs Bradley'; *ibagam, ta kelnátenna deyta nateŋ* 'Tell her to steam the vegetables'.

Similarly, comparable to oblique object attributes are clauses headed by *idi* 'when (in the past)': *napan idyay filipínas, idi madama ti gubat* 'He went to the Philippines while the war was going on'.

Proclitic *nu* makes clauses of unfulfilled event and question clauses: *nu tulúŋannak, bilsígek deyta káyu* 'If you help me, I'll chop this wood'; *dyak mamáti nu mabalin* 'I don't think it is possible'; *dyak makíta nu ápay ŋa kitkitáʔen ni jeni bradli únay-únay dagiti ubbiŋ* 'I can't see why (*ápay*, conjunct) Jennie Bradley keeps bothering so much with the children'.

tapnu 'so that'; *aláʔem deytuy, tapnu adda gastwem* 'Take this so you'll have something to spend'.

32. An important proclitic, joined in close juncture, is *kas* 'like, as'. It rarely occurs without other particles: *nu kaspagarigan makauŋet ni ámutayu* 'if, just by way of supposition, our master gets angry'. It is often combined with *ta* and *tuy*, which here are used non-proclitically: *ó, kasta!* 'Oh, it's like that, is it?' *sápay kastuy?* 'Why is it like this?' *maŋŋeg ti adú na parparutuput, kastá met ti naragsak ŋa katkatáwa dagiti adú ŋa ubbiŋ* 'One can hear many reports of firecrackers, as well as the gay laughter of many children'. Otherwise *kas* is usually combined with enclitic *la* 'only': *kasla matay* 'He is as if dead'; *sumrek ŋa kasla áŋin* 'She comes in like a gust of wind'; *kaskantula kankanáyun ŋa nasadut* 'You will always seem to be lazy'; here *kas* has attracted the enclitics *ka* and *tu*.

33. Various phrases are thus made of two or more particles. Thus *wen* 'yes' with *nu* means 'or': *ti kawáyan wen nu ti búlu* 'ordinary bamboo or bulu bamboo'. *intu nu* 'when' of future occurrence in certain combinations: *intunu bigat* 'tomorrow', *intun umay ŋa tawen* 'in the coming year; next year'. The following illustrate some other combinations: *dyak mapmapan, nú ket di makaŋŋégak iti dákes* 'I don't go there, because I might hear bad things'; *intú ket dí nu makakítaʔak iti dákes* 'I might see evil things'; *itedku ti biŋlaymu, nulá ket di ikkannak iti bibiŋka* 'I'll give (you) your share only if you give me some of the pancake'.

34. Other heads of phrases are certain full words. Such are *gapu* 'cause' with *ta*: *napalálu ti uŋet ni mister bradli, gaput ta dagiti amerikánus, pinalpaltúganda dagiti filipínus* 'Mr Bradley's anger grew great, because the Americans were shooting at the Filipinos'; *aggapu* '(comes) from': *sumrek ni súsi aggapu iti kumedor* 'Susie enters from the dining room' (but also: *ni súsi, sumrek ŋa aggapu iti hardin* 'Susie comes in from the garden'); *aginga* 'ends' with *na* 'of-it' and *di* 'not': *agtaráyak agingánat dínak pasardeŋen*

'I'll run until you stop me'; *agálu* 'gets big; especially': *kunáda ŋa ti meysa ŋa táʔu, nu matay, aglálu nu adú ti nabáti ŋa duŋduŋwenna, kasakbáyan ti ipapán ti kararwána idyay sabáli ŋa lúbuŋ, umayna kitáʔen, nu naistílu met láʔeŋ ti byag dagiti nabáti* 'They say that when a person dies, especially if the loved ones left behind are many, before the going of his soul to the other world, he comes to see whether the life of those left behind is (reasonably, fairly, *met láʔeŋ*) good'; this illustrates also *kasakbáyan* 'before'; similarly *kalpasan* 'after'; *paŋgep*, *maipaŋgep* 'on account of, about': *adda maarámidku paŋgep ti uppat ti hulyu* 'I have things to do in connection with the Fourth of July'.

35. Paratactic and parenthetic expressions are fairly common and well marked by phrase finals: *adda meysa ŋa paduŋ agnaʔed iti ígid ti karayan* 'Once upon a time there was a turtle, (and he) lived by the bank of a river'.

Names and other terms of address are in open form, and often appear as enclitics, without phrase finals: *anakku, díka agarʔarámid iti dákes* 'My child, don't do bad things'; *ápay paduŋ ŋa agibʔibitka?* 'Why, Turtle, art thou weeping?'; *ápay baruk saʔanmú ŋa inbaga kadakami ti gandatmu?* 'Why, my boy, haven't you told us (of) your intention?'

36. Minor sentences are few, since in general a full word constitutes a full sentence (para. 10, p. 96): *wen* 'yes'; *saʔan* 'no'; *anay!* 'ouch!'; *salʔit!* 'lightning; Confound it!'

7 B. L. Whorf

Grammatical Categories[1]

B. L. Whorf, 'Grammatical categories', *Language*, vol. 21, 1945, pp. 1–11.

The very natural tendency to use terms derived from traditional grammar, like verb, noun, adjective, passive voice, in describing languages outside of Indo-European, is fraught with grave possibilities of misunderstanding. At the same time it is desirable to define these terms in such a way that we can avail ourselves of their great convenience and where possible apply them to exotic languages in a scientific and consistent way. To do this, we must re-examine the types of grammatical category that are found in languages, using a world-wide view of linguistic phenomena, frame concepts more or less new, and make needed additions to terminology. These observations apply *pari passu* to English, which, hardly less than some American Indian languages, is off the pattern of general Indo-European.[2]

In the reaction from conventional grammars of American languages based on classical models, there has been a tendency to restrict attention to the morphemes by which many grammatical forms are marked. This view loses sight of various word-classes that are marked not by morphemic tags but by types of patterning, e.g. by the systematic avoidance of certain morphemes, by lexical selection, by word-order that is also CLASS-ORDER; in general by association with definite linguistic configurations. At the beginning of investigation of a language, the 'functional' type of definition, e.g. that a word of a certain class, say a 'noun', is 'a word which does so-and-so', is to be avoided when this is the ONLY test of distinction applied; for people's conceptions of what a given word 'does' in an unfamiliar language may be as diverse as their own native languages, linguistic educations, and philosophical predilections. The categories studied in grammar are those recognizable through facts of a configurational sort, and these facts are the same for all observers. Yet I do not share the complete distrust of all functional definitions which a few modern grammarians seem to show. After categories have been outlined according to configurative facts, it may be desirable to employ functional or operational symbolism as the investigation proceeds. Linked with configurative data, operational descriptions become valid as possible ways of stating the MEANING of the forms, 'meaning' in such cases being a characterization which succinctly accounts for all the semantic and configurational facts, known or predictable.

---

[1] This paper was written late in 1937 at the request of Franz Boas, then editor of the *International Journal of American Linguistics*. The manuscript was found in the Boas collection by C. F. Voegelin and Z. S. Harris. The author died on 26 July 1941.

[2] The author wishes to acknowledge his indebtedness to his colleagues, Dr George L. Trager and Dr Morris Swadesh, with whom some of these questions of category have been discussed.

We may first distinguish between OVERT CATEGORIES and COVERT CATEGORIES.

An overt category is a category having a formal mark which is present (with only infrequent exceptions) in every sentence containing a member of the category. The mark need not be part of the same word to which the category may be said to be attached in a paradigmatic sense, i.e. it need not be a suffix, prefix, vowel-change, or other 'inflection', but may be a detached word or a certain patterning of the whole sentence. Thus in English the plural of nouns is an overt category, marked usually in the paradigm word (the noun in question) by the suffix *-s* or a vowel-change, but in the case of words like *fish*, *sheep* and certain gentilic plurals, marked by the form of the verb, the manner of use of the articles, etc. In *fish appeared* the absence of any article denotes plural, in *the fish will be plentiful* a pluralizing adjective denotes it, in *the Chinese arrived* and *the Kwakiutl arrived*, the definite article coupled with lack of a singular marker like *person*, *Chinaman*, or *Indian* denotes plural. In all these cases plural is overtly marked, and so with few exceptions are all noun plurals in English, so that noun-plural is an overt category in English.[3] In Southern Paiute the subject-person of a verb is marked by a sub-lexical element (or 'bound morpheme') that cannot stand alone, like Eng. *-s*; but it need not be attached to the verb, it may be attached to the first important word of the sentence. In English what may be called the potential mode of the verb is an overt category marked by the morpheme *can* or *could*, a word separate in the sentence from the verb but appearing in every sentence containing the category. This category is as much a part of the verb system of morphology as though it were denoted by a bound element in a synthetic Algonkian or Sanskrit verb; its morpheme *can* may replace coordinate elements in the same modal system, e.g. *may*, *will*, but it may not, like a mere lexical item (e.g. *possibly*) be simply added to them. In Hopi also there is a rigid system of mutually exclusive 'modalities' denoted by detached words.

---

[3] There is of course a minority group of possible or theoretically possible sentences, e.g. *The fish appeared*, in which plural is not distinguished from singular. But in actual speech such sentences are embedded in a larger context which has already established the plurality or the singularity of the thing discussed. (Otherwise such a sentence is not likely to occur.) Such minority types are not considered in the distinction between overt and covert, i.e. they do not prevent a category from being classed as overt. In covert categories the unmarked forms are relatively numerous, often in the majority, and are undistinguished even by context.

A covert category is marked, whether morphemically or by sentence-pattern, only in certain types of sentence and not in every sentence in which a word or element belonging to the category occurs. The class-membership of the word is not apparent until there is a question of using it or referring to it in one of these special types of sentence, and then we find that this word belongs to a class requiring some sort of distinctive treatment, which may even be the negative treatment of excluding that type of sentence. This distinctive treatment we may call the REACTANCE of the category. In English, intransitive verbs form a covert category marked by lack of the passive participle and the passive and causative voices; we cannot substitute a verb of this class (e.g. *go*, *lie*, *sit*, *rise*, *gleam*, *sleep*, *arrive*, *appear*, *rejoice*) into such sentences as *It was cooked*, *It was being cooked*, *I had it cooked to order*. An intransitive thus configuratively defined is quite a different thing from the 'dummy' intransitive used in traditional English grammar; it is a true grammatical class marked by these and other constant grammatical features, such as non-occurrence of nouns or pronouns after the verb; one does not say *I gleamed it*, *I appeared the table*. Of course compound formations involving these same lexemes may be transitive, e.g. *sleep* (*it*) *off*, *go* (*him*) *one better*. In the American colloquial forms *go haywire*, *go South Sea Islander*, etc., the word or phrase after the verb is a covert adjective, cf. *go completely haywire*.

Another type of covert category is represented by English gender. Each common noun and personal given name belongs to a certain gender class, but a characteristic overt mark appears only when there is occasion to refer to the noun by a personal pronoun in the singular number – or in the case of the neuter it may be marked by the interrogative and relative pronouns *what*, *which*. The grammatical alignment is no less strict than in an overt gender system like that of Latin, where most nouns bear their gender mark. No doubt for many English common nouns a knowledge of actual sex and of scientific biological and physical classification of objects could serve a foreigner in lieu of knowledge of the grammatical classes themselves, but such knowledge would be of only limited use after all, for the greater part of the masculine and feminine classes consists of thousands of personal names, and a foreigner who knows nothing of the cultural background of Western European Christian names must simply learn, i.e. observe, that *Jane* belongs to the 'she' group and *John* to the 'he' group. There are plenty of names of overt similarity but contrasted gender, e.g. *Alice*: *Ellis*, *Alison*: *Addison*, *Audrey*: *Aubrey*, *Winifred*: *Wilfred*, *Myra*: *Ira*, *Esther*:

Lester.[4] Nor would knowledge of any 'natural' properties tell our observer that the names of biological classes themselves (e.g. *animal*, *bird*, *fish*, etc.) are 'it'; that smaller animals usually are 'it'; larger animals often 'he'; dogs, eagles and turkeys usually 'he'; cats and wrens usually 'she'; body-parts and the whole botanical world 'it'; countries and states as fictive persons (but not as localities) 'she'; cities, societies, and corporations as fictive persons 'it'; the human body 'it'; a ghost 'it'; nature 'she'; watercraft with sail or power and named small craft 'she'; unnamed row-boats, canoes, rafts 'it', etc. The mistakes in English gender made by learners of the language, including those whose own languages are without gender, would alone show that we have here covert grammatical categories, and not reflections in speech of natural and non-cultural differences.

The classes of nouns based actually or ostensibly upon shape, in various American languages, may be either overt or covert. In Navaho they are covert. Some terms belong to the round (or roundish) class, others to the long-object class, others fall into classes not dependent on shape. No overt mark designates the class in every sentence. The class mark as in English gender is a reactance; not a pronoun, however, but a choice between certain verb stems that go definitely with one class and no other, although there are very many verb stems indifferent to this distinction. I doubt that such distinctions, at least in Navaho, are simply linguistic recognitions of non-linguistic, objective differences that would be the same for all observers, any more than the English genders are; they seem rather to be covert grammatical categories. Thus one must learn as a part of learning Navaho that 'sorrow' belongs in the 'round' class. One's first and 'common sense' impression of covert categories like English gender and Navaho shape-class is that they are simply distinctions between different kinds of experience or knowledge; that we say *Jane went to her house* because we know that Jane is a female. Actually we need not know anything about Jane, Jane may be a mere name; yet having heard this name, perhaps over the telephone, we say *What about her?* Common sense may then retreat a step further and say that we know the name Jane to be given only to females. But such experience is linguistic; it is learning English by observation. Moreover it is easy to show that the pronoun agrees with the name only, not with the experience. I can bestow the name *Jane* on an automobile, a skeleton, or a cannon, and it will still require *she* in pronominal references. I have two goldfish; I name one *Jane* and one *Dick*. I can still say *Each goldfish likes its food*, but not *Jane likes its food better than Dick*. I must say *Jane likes her*

---

[4] There are a very few names of indeterminate or double gender: *Frances (Francis)*, *Jessie (Jesse)* or *Jess*, *Jean (Gene)*, *Jocelyn*, *Sidney*, *Wynne*, and perhaps a few others. The number is increased if we include nicknames like *Bobby*, *Jerry*, etc.; but all in all such cases are relatively so few that they in no way disturb our alignment of facts.

food. The word *dog* belongs to a common gender class with a preference for *he* and *it*, but the gender-classed given name of a dog determines its own pronoun; we do not say *Tom came out of its kennel*, but *Tom came out of his kennel*, *Lady came out of her kennel*, *The female dog came out of its* (or *her*) *kennel*. 'Doggish' names like *Fido* are of the 'he' class: *Towser came out of his kennel*. We say *See the cat chase her tail*, but never *See Dick chase her tail*. The words *child*, *baby*, *infant* belong to the common class and can take *it*, but the given names of children take either *he* or *she*. I can say *My baby enjoys its food*, but it would be linguistically wrong to say *My baby's name is Helen – see how Helen enjoys its food*. Nor can I say *My little daughter enjoys its food*, for *daughter*, unlike *baby*, is grammatically in the feminine class.

Likewise with various covert categories of exotic languages; where they have been thought to be recognitions of objective differences, it may rather be that they are grammatical categories that merely accord up to a certain point with objective experience. They may represent experience, it is true, but experience seen in terms of a definite linguistic scheme, not experience that is the same for all observers. On the other hand, the distinctions between present and absent, visible and invisible, made in many American languages, may well represent experiential differences; and again we may have such experiential differences engrafted upon purely grammatical classifications, yielding mixed classes such as 'experiential-present plus grammatical-feminine'.

A covert category may also be termed a CRYPTOTYPE, a name which calls attention to the rather hidden, cryptic nature of such word-groups, especially when they are not strongly contrasted in idea, nor marked by frequently occurring reactances such as pronouns. They easily escape notice and may be hard to define, and yet may have profound influence on linguistic behavior. The English intransitive verbs as configuratively defined above are a cryptotype. A similar cryptotype comprises the verbs of 'copulative resolution' (*be*, *become*, *seem*, *stay*, *remain*, etc.), which also lack the passive and causative but may be followed by nouns, pronouns, and adjectives. Transitives (a cryptotype which includes *run*, *walk*, *return*, etc. – indeed most English verbs) possess the passive and causative and may be followed by nouns and pronouns but not by adjectives alone. Names of countries and cities in English form a cryptotype with the reactance that they are not referred to by personal pronouns as object of the prepositions *in*, *at*, *to*, *from*. We can say *I live in Boston* but not *That's Boston – I live in it*. A word of this cryptotype is referred to by *there* or *here* in place of *in it*, *at it*, *to it*, and by *from there* (*here*) in place of *from it*. In various American languages such place-names constitute a grammatical class; in Hopi they lack the nominative and objective cases, occurring only in locational cases; in Aztec they bear characteristic endings and exclude the use of certain prepositions.

English adjectives form two main cryptotypes with sub-classes. A group referring to 'inherent' qualities – including color, material, physical state (solid, liquid, porous, hard, etc.), provenience, breed, nationality, function, use – has the reactance of being placed nearer the noun than the other group, which we may call one of non-inherent qualities, though it is rather the residuum outside the first group – including adjectives of size, shape, position, evaluation (ethical, esthetic, or economic). These come before the inherent group, e.g. *large red house* (not *red large house*), *steep rocky hill*, *nice smooth floor*. The order may be reversed to make a balanced contrast, but only by changing the normal stress pattern, and the form is at once sensed as being reversed and peculiar. The normal pattern has primary stress either on the noun *(steep rocky hi′ll)* or on the inherent adjective *(pretty Fre′nch girl)*. We cannot simply reverse the order of adjectives and say *French pre′tty girl* – the form suggests a contrasted *French pla′in girl* but the pattern of so contrasting adjectives is un-English; the proper contrast is *pla′in French girl*. We can however reverse the adjectives by altering the stress pattern and say *Fre′nch pretty girl*, if in contrast with e.g. *Spa′nish pretty girl*, though such forms are clearly exceptional.

The contrasting term PHENOTYPE may be applied to the overt category and, when no ambiguity results, to the mark which accompanies the overt category in the sentence.

The distinction between overt and covert categories, or PHENOTYPES and CRYPTOTYPES, is one of two distinctions of supreme importance in the theory of grammatical categories. The other is the distinction between what may be called SELECTIVE CATEGORIES and MODULUS CATEGORIES.

A selective category is a grammatical class with membership fixed, and limited as compared with some larger class. A PRIMARY selective category, or LEXEMIC category, is one compared to which the next larger class is the total lexicon of the language. Certain semantic and grammatical properties are assured in the word by selecting it from a certain class of fixed membership not coterminous with the whole vocabulary. In order that a certain grammatical property may be 'in the lexeme' it cannot be in all lexemes. The familiar 'parts of speech' of most European languages, but not of English, are lexemic categories. The situation in English is peculiar, and will be touched upon later. Lexemic categories may be either overt or covert. Hopi is an example of a language in which they are covert. Possibly Maya may be another such case, though we lack clear information on that point. In Hopi there is no distinction in the simplex (bare stem) forms between nouns and verbs, and sentences are possible in which there is no distinction in the sentence. Thus *le·'na* or *pam le·'na* means 'it is a flute' and *pe·'na* or *pam pe·'na* means 'he writes it'. Hence nouns and verbs MAY BE alike in overt characteristics. But it is easily possible to make sentences in which *le·'na* appears with case suffixes and in other forms quite impossible for *pe·'na*, and vice versa. One has to learn, and cannot always tell from the sentence, that *le·'na* and *pe·'na* belong to different compartments of the lexicon.

It is probably more common to find lexemic classes that are overt, as in Latin, French, Aztec, Tübatulabal, Taos, and Navaho. In French, *ange* and *mange* belong to different compartments of the vocabulary (noun and verb) and there is always a feature in the sentence that tells which; one does not find such pairs as *il mange*: *il ange*, *c'est un ange*: *c'est un mange*. It may be possble to have *Ange!* versus *Mange!*, but special and abbreviated types of sentence like these with their lack of formal distinctions do not justify calling the categories covert. In Latin, Aztec, Tübatulabal and Taos, the distinction is marked not only in the sentence, but usually in the paradigm word itself. Yet this overt mark of the noun, verb, or other 'part of speech' cannot usually be transferred to a lexeme outside of the proper group. The mark that goes with a covert lexemic class need not stand for any other category such as case, person, or tense, though it does e.g. in Latin, Greek, and Sanskrit. The 'absolutive suffixes' found attached to lexemic nouns in most Uto-Aztecan languages have basically no other character than that of such class-marks, though in Aztec they are also tied up with number; and needless ingenuity has been wasted in trying to make them out to be 'articles' or the like. The absolutive suffixes in Taos go with the selective class of nouns but indicate gender and number also. In Latin the distinction between the nouns (including in this class the adjectives) and the verbs is selective and overt, but that between adjectives and substantives is selective and covert; compare *est gladius* and *est bonus*. As with all covert classes, the distinction is revealed upon forming the proper type of sentence; *est bona* occurs, but not *est gladia*.

Lexemic categories include not only nouns, verbs, adjectives and other 'parts of speech', but also 'full' words and 'empty'[5] words or stems, as in Chinese and perhaps the Wakashan languages, and still other types of distinction; e.g. in Algonkian the lexemic classes include large groups of stems having different combinatory powers and different positions in the verbal complex.

---

[5] An 'empty' word or stem is probably one that is highly specialized for grammatical or syntactic indication, perhaps in a way that does not admit of being assigned a concrete meaning. For example, such a form might have no other meaning than to serve as the reactance of some other category, or as the signature of a modulus category (see the next paragraph).

A modulus category is a non-selective category, i.e. it is generally applicable and removable at will. Depending on its type it may be applied either to any 'major word' (any word excepting small and specialized selective classes, e.g. 'particles'), or, more often, to any word coming within a certain pre-requisite larger category, which may be either selective or another modulus category. The cases, tenses, aspects, modes and voices of Indo-European and Azteco-Tanoan languages are modulus categories, applicable at will to words belonging to the proper larger category – cases being moduli of the larger category of nouns; aspects, tenses, etc. moduli of the larger category of verbs (Whorf and Trager, 1937). Hence the person versed only in Indo-European types of grammar poses to himself the distinction between selective and modulus classes (or between selectivity and modulation) as the distinction between 'parts of speech' on the one hand and 'grammatical forms' of the aspect, tense, and voice type on the other. But in widely different types of speech these familiar types of meaning and function cease to be associated with selectivity and modulation in the same way; entirely different alignments there hold sway in the grammar, and until this is recognized an adequate conception of the grammar cannot be obtained. It is not necessary to have large categories, such as nouns and verbs, in order to have such modulus categories as aspect. In Nitinat (and presumably in the closely related Nootka and Kwakiutl) all major words have aspects, such as durative, momentaneous, inceptive, etc. – both the word for 'run' and the word for 'house' always bear some element marking this aspect (Swadesh and Swadesh, 1933).

We may use the term modulus alone to denote the distinctive class meaning and function of the category; thus the present-participle meaning is a modulus in English. We may also use modulus to mean the grammatical operation of producing one such meaning, and hence, where no ambiguity results, to mean the element or pattern that marks the modulus. Thus we can say that in English the present-participle modulus is the suffixing of *-ing*, or for short that it is *-ing*. Where greater preciseness is desirable, we may call the overt mark the (or a) SIGNATURE of the modulus. This distinction is ultimately important; sometimes it is necessary to distinguish several signatures of the same modulus. In illustrating overt categories we cited the English noun-plural, which is a modulus category. The modulus, or plural type of meaning, is one and the same thing throughout the various examples; but the signatures whereby this plural modulus may be applied to the word *fish* are different from one example to the next. To these signatures we may add *-s* or *-es*, giving *fishes*. Since *sheep*, *deer*, *moose*, *caribou*, etc. belong to a cryptotype that excludes *-s*, and 'fishermen's fish' such as *trout*, *bass*, *salmon*, *mackerel*, *cod*, etc. (contrasted with 'low-grade fish', e.g. *sharks*, *skates*, *eels*, *sculpins*, etc.) belong to another such cryptotype, we cannot use this last signature for them. As this example shows, it is not necessary to have one-to-one correspondence between moduli and signatures. Where a high degree of such one-to-one correspondence obtains it has often been the custom to apply the graphic but not very scientific catchword 'agglutinative' to the language. Languages of the typical 'agglutinative' type, such as Turkish, have been referred to as if they had such one-to-one correspondence, and moreover as if they had no categories but modulus categories. The grammar of Yana (Hokan stock, California) consists largely of moduli, but has also a few selective categories, e.g. a class of stems which must stand first in the verbal complex and a class which must stand second.

A distinction of the same semantic type as that between verbs and nouns in selective categories may be handled by modulus categories instead. That is, the possible moduli include not only voice, aspect, etc., but also VERBATION and STATIVATION.[6] Whenever, as e.g. in Yana, the mere application of certain distinctive suffixes or other signatures makes a 'verb' out of any stem, then we do not have a class of verbs in the same sense as in French, Latin, Greek, Hopi, Aztec, Taos and Navaho, i.e. a selective class. We have verbations instead of such verbs. The so-called verbs and nouns of Semitic are moduli, applicable to lexemes in general by signatures consisting largely of vowel-consonant sequence patterns, though there may be occasional gaps in the universality of lexical applicability. In Hebrew we have *e-e* as one of several signatures for stativation and *ā-a* as one of several for verbation, e.g. *berek* 'knee':*bārak* 'he kneeled', *derek* 'road': *dārak* 'he marched', *geber* 'man, as virile or strong':*gābar* 'he was strong', *hebel* 'cord':*hābal* 'he bound', *melek* 'king':*mālak* 'he reigned', *qedem* 'antecedence':*qādam* 'he was before', *regel* 'foot':*rāgal* 'he went on foot'.[7] There are no doubt many Hebrew 'nouns' for which we do not know the verbation form in texts, but this seems to be largely because the textual Hebrew that we know does not represent the full resources of the ancient living language; Arabic shows better the general applicability of these moduli to the great majority of lexemes. But verbs and nouns which are modulus categories may be found nearer home than Semitic. The lexicon of English contains two major selective divisions. One division, consisting mostly of long words and words with certain endings, contains selective verbs like *reduce*, *survive*, *undertake*, *perplex*, *magnify*, *reciprocate*, and selective nouns like *instrument*, *elephant*, *longevity*, *altruism*. A limited number of short words belong also to the group of selective nouns and verbs, e.g. *heart*, *boy*, *street*, *road*, *town; sit*, *see*, *hear*, *think*. In this selective vocabulary English is like French or Hopi. The other part of the lexicon, mostly the shorter words but some long ones, contains bare lexemes to which either verbation or stativation may be applied at will, e.g. *head*, *hand*, *stand*, *walk*, *exchange*, *sight*, *skin*, *weave*, *dog*, *surrender*, *massage*, etc.[8] This part of the vocabulary is like Arabic, though the signatures are of a quite different sort. Those for stativation include the articles, plural signatures, position after possessive pronouns and selective adjectives; those for verbation include position after a nominative pronoun, position before a pronoun, noun, or stativation, the tense forms, the verbal auxiliaries and modal particles, etc.

---

[6] Stativation is a term used to denote the modulus of forms which are contrasted with verbations in a way similar to that in which nouns, as a selective category, are contrasted with verbs in the languages that have such a contrast. It is used here instead of 'nomination' or 'nominalization' because these terms through past usage have come to suggest derivations rather than moduli, while 'stativation' helps us to think of the form not as a noun derived from a verb, but simply as a lexeme which has been affected by a certain meaningful grammatical coloring as a part of certain configurations.

[7] Since these Hebrew examples are used only to illustrate vowel-patterns, they are written in approximate morphophonemic orthography, which does not attempt to show the distinction between the stops *b*, *g*, *k*, etc. and the spirants which replace them after vowels under regular statable conditions.

There may be wide variability in the semantic relations between verbations and stativations in the same language. When contrasted with the corresponding stativations, verbations may seem to add in an inconstant manner such ideas as 'he engaged in' *(hunt*, *jump*, *dance)*, 'behave like' *(mother*, *carpenter*, *dog)*, 'be in' *(lodge*, *hive)*, 'put in' *(place*, *seat*, *pocket*, *garage)*, 'make, add, install' *(weave*, *plant*, *roof*, *pipe*, *tin)*, 'take away' *(skin*, *peel*, *husk*, *bone)*, 'get' *(fish*, *mouse)*, 'use' *(spear*, *hammer*, *fiddle*, *bugle)*; while on the other hand stativations seem to add inconstantly such ideas as 'result' *(weave*, *plant*, *form)*, 'means' *(paint*, *trail)*, 'action or place' *(walk*, *slide*, *step*, *drop)*, 'instrument' *(lift*, *cover*, *clasp*, *clip)*, etc. This inconstancy, or better elasticity, in certain aspects of the meaning, seen in Semitic as well as English, is characteristic of the simple moduli of verbation and stativation, and it may be contrasted with the condition of having a number of different moduli, each a different specialized type of verbation or stativation, which appears to be the situation in Alaskan Eskimo. It merely means that in a language with simple primary types of moduli the meaning of the individual lexeme is more or less under the sway of the entire sentence, and at the mercy of the manifold potentialities of connotation and suggestion which thereby arise.

Can there be languages not only without selective nouns and verbs, but even without stativations and verbations? Certainly. The power of making predications or declarative sentences and of taking on such moduli as voice, aspect, and tense, may be a property of every major word, without the addition of a preparatory modulus. This seems to be the case in Nitinat and the other Wakashan languages. An isolated word is a sentence; a sequence of such sentence words is like a compound sentence. We might ape such a compound sentence in English, e.g. 'There is one who is a man who is yonder who does running which traverses-it which is a street which elongates', though the exotic sentence consists simply of the predicative lexemes 'one', 'man', 'yonder', 'run', 'traverse', 'street', and 'long', and the proper translation is 'A man yonder is running down the long street'. Such a structure might or might not be found in an isolating language; again it might or might not be found in a polysynthetic one like Nitinat. The polysynthetic language might or might not fuse some of the lexemes into long synthetic words, but it would doubtless have the power in any case of fusing in a great many aspectual, modal and connective elements (signatures of moduli). Of such a polysynthetic tongue it is sometimes said that all the words are verbs, or again that all the words are nouns with verb-forming elements added. Actually the terms verb and noun in such a language are meaningless. The situation therein is radically different from e.g. Hopi, for though in the latter *le·'na* 'it is a flute' and *pe·'na* 'he writes it' are both complete sentences, they are words which are not equally predicative in all positions of a sentence, and they also belong to selective covert classes of noun and verb that in general take different inflections, and look alike only in particular types of sentence. In Hopi the verb-noun distinction is important on a selective basis; in English it is important on a modulus basis; in Nitinat it seems not to exist.

---

[8] Adjectivation in English is another modulus which is applied both to bare lexemes and to selective nouns, but there are also selective adjectives, and these are not modulated into substantives.

So far we have dealt with categories which are distinct both configuratively and semantically, and these are the typical formulations of grammar. But we also have word groups which are configuratively distinct yet have no difference in meaning; these we may call ISOSEMANTIC or purely formal classes. They in turn are of two sorts corresponding to selective and modulus in the semantic categories, but here better styled SELECTIVE and ALTERNATIVE. Selective isosemantic classes are typified by 'declensions' and 'conjugations', those very common features of languages the world over; richly developed in Latin, Sanskrit, Hopi and Maya, less developed in Semitic, English ('strong' and 'weak' verbs) and Aztec, and almost lacking in Southern Paiute. They also include gender-like classes without semantic difference, as in Bantu and in certain of the genders of Taos (all these might be called 'declensions' with pronominal agreement or the like); classes requiring different position in a sentence or complex without difference in type of meaning (stem-position classes in Algonkian); and classes requiring different signatures for the same modulus without difference in type of meaning, e.g. in Hebrew the segholate *(e-e)* 'nouns' and parallel stativation-groups. Alternative isosemantic classes are what their name implies, e.g. the English group comprising *don't*, *won't*, *shan't*, *can't*, etc. and the group of *do not*, *will not*, *shall not*, *cannot*. In this case we could perhaps speak of a modulus of brevity, convenience, or colloquial attitude which is applied in the former group. Alternative classes sometimes show STYLISTIC as opposed to grammatical difference. In other cases there seems to be no generalizable difference, as in English *electrical*, *cubical*, *cyclical*, *historical*, *geometrical* versus *electric*, *cubic*, *cyclic*, *historic*, *geometric*.

There remains another type of distinction: SPECIFIC CATEGORIES and GENERIC ones. A specific category is an individual class existing in an individual language, e.g. English passive voice, Hopi segmentative aspect. A generic category, in the restricted sense of application to a particular language, is a hierarchy formed by grouping classes of similar or (and) complementary types, e.g. case in Latin, voice in Hopi. Here much depends on both the insight and the predilections of the systematizer or grammarian, for it may be easy to build up specific categories into very logical schemes, yet what is rather desired is that such generic categories should represent systems which the language itself contains. We do well to be skeptical of a grammarian's systematization when it is full of ENANTIOMORPHISM, the pairing with every category of an opposite which is merely the lack of it. Specific categories of seemingly opposite meaning such as passive voice and active voice (when this term 'active' means merely 'non-passive') should be brought into one generic category ('voice') only when they are more than two, or when, if only two, taken together they contrast as a unit with some other system of forms.

Finally, in a still wider sense generic categories may be so formulated as to become equivalent to the concepts of a general science of grammar. Such categories are made by grouping what seem to us to be SIMILAR SPECIFIC CATEGORIES IN DIFFERENT LANGUAGES. Only in such a sense can we speak of a category of 'passive voice' which would embrace the forms called by that name in English, Latin, Aztec and other tongues. Such categories or concepts we may call TAXONOMIC categories, as opposed to DESCRIPTIVE categories. Taxonomic categories may be of the first degree, e.g. passive voice, objective case; or of the second degree, e.g. voice, case. Perhaps those of the second degree are the more important and ultimately the more valuable as linguistic concepts, as generalizations of the largest systemic formations and outlines found in language when language is considered and described in terms of the whole human species.

References

SWADESH, M. H., and SWADESH, M. (1933), 'A visit to the other world: a Nitinat text', *Int. J. Amer. Linguistics*, vol. 7, pp. 34.

WHORF, B. L., and TRAGER, G. C. (1937), 'The relationship of Uto-Aztecan and Tanoan', *Amer. Anthrop.*, vol. 39, pp. 609–24.

8 V. Yngve

The Depth Hypothesis

V. Yngve, 'The depth hypothesis', *PAMS*, 1961, vol. 12, pp. 130–8.

It is well known that there are striking contrasts as well as striking similarities between a natural language like English, and the languages, or notation systems, of mathematics. We have, on the one hand, the wealth of complexity of natural languages, particularly on the level of sentence structure, and, on the other hand, the simplicity and elegance of the mathematical notations. Some of the syntactic features found in English have the obvious function of role-marking and subordination, but it has not been clear why there is such a diversity of role-marking and subordinating devices in view of the fact that a single device can be used, as in the Łukasiewicz or Polish prefix notation. Then, there is a long list of other complications in English syntax whose functions have been completely obscure. The many types of discontinuous constituents furnish examples of these syntactic devices with no obvious function.

In this paper, a 'depth hypothesis' will be proposed as an explanation for many of these previously unexplained features of English syntax.[1] Psychologists have measured what they call the span of immediate memory. We are able to memorize at a glance and repeat back correctly about seven random digits, about seven nonsense words, about seven items. This has been known for a long time, and has been summarized by George Miller of Harvard (1956). Apparently we have to live with this restriction when we talk. We can only remember about seven grammatical or syntactic constraints at one time. The depth hypothesis states that much of the syntactic complexity of English can be understood in terms of this memory restriction. The syntax of English contains many devices for automatically keeping utterances within the bounds of this restriction, and it contains many devices for effectively circumventing the restriction so as to regain the loss of expressive power that this restriction on immediate memory span would imply. The depth hypothesis predicts that *all* languages have extensive syntactic complications for the same purpose. This is subject to test.

A restriction on immediate memory span would not, however, be expected to have any influence on the structure of mathematical notations because they are written notations.

---

[1] A more detailed and complete presentation has been published in Yngve (1960).

The way in which a limited memory can affect the syntactic structure of a language can be understood on the basis of a simple model of sentence production. This model can be easily programmed on a computer. It was, in fact, designed for use in a mechanical translation system. The linguistically important features of the model can be understood on the basis of two assumptions. The first is that a phrase-structure or immediate constituent framework will be used for the description of the syntax of the language. But since such a description is static, it is necessary to introduce the aspect of time – the order in which the model does things – by a second assumption: that the words of a sentence are produced one at a time in their natural time sequence – that is, left-to-right according to conventional orthography – and that phrase-structure rules are applied by expansion – that is, from the top, down, in our diagrams.

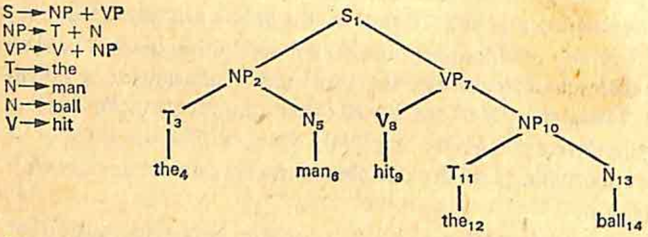


Figure 1 The production of a sentence involves the application of grammar rules like those shown on the left. The rules are applied by expansion from left to right. The subscript numbers show the order in which the rules have been applied

Figure 1 shows how these assumptions would work out in the actual production of sentences by the model. Somewhere in the memory of the device is stored a set of phrase-structure rules, like those listed on the left, that give the immediate constituents for every construction in the language. It is assumed that this list of rules is finite and that there is no linguistic significance to the order in which the rules are listed. Any order will do – an alphabetic order may be convenient. At the top of the tree is a symbol S for sentence. The first rule to be applied says to expand the symbol S into its immediate constituents: a subject noun phrase and a predicate verb phrase. Then the subject noun phrase is expanded into an article and a noun. And so on. The order in which the constructions are expanded into their constituents is given by the subscript numbers. This order involves always expanding the left-hand member of every construction first, and when reaching the end of a branch, retracing to the next higher right-hand member. In this way the words are produced in their proper left-to-right order.

Although the grammar consists of a finite number of rules, it is a property of the device that it can produce any sentence chosen from an infinite set of sentences. It can do this by reapplying rules indefinitely to provide clauses within clauses within clauses. Figure 2 illustrates how the same rule can be applied twice. At the top, S is expanded into NP and VP, and at the right S is again expanded into NP and VP, in a dependent clause. This can go on without limit. 'The girl knows that the papers reported that the woman said that the man', and so on.

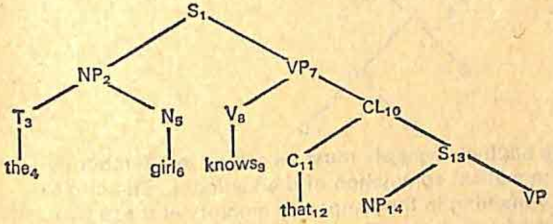


Figure 2 Indefinitely long sentences can be produced by reapplying certain rules. Here the rule S→NP+VP has been applied at node 1 and again at node 13

The way in which a limited span of immediate memory comes into the picture can be seen from the following considerations. In producing a sentence, the device must remember somehow what it is committed to do next by the rules of the language. Having expanded S into a subject NP and a predicate VP, it goes on to expand the left-most constituent, NP. But somewhere in its memory it has to store the information that when it gets finished with all the branches hanging onto the NP, it is then committed to expand a VP, otherwise it will not have a grammatical English sentence. Similarly, after expanding NP into T and N, it must store away the N while it is expanding the T. In this way, every time the device goes down a left-hand branch, it must store in its temporary memory one symbol for each step taken down the branch. In the left diagram of Figure 3, the machine is working on node 5 of a sentence. It has stored away four node names, represented by the circles. In the diagram on the right, this same structure is shown at a later stage. The machine has finished with node 5, obtained nodes 6 and 7 from its memory and expanded them, and is now working on node 11. It still has four symbols stored away in its memory: two of the original ones and two new ones that resulted from the expansion of nodes 7 and 10. It will eventually have to return and expand them all.

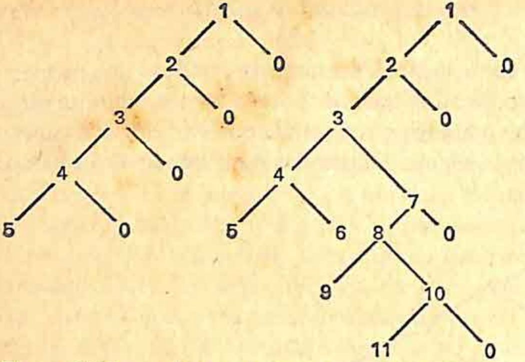


Figure 3 When rules are applied, symbols must be stored in a temporary memory to assure a grammatical completion of the sentence. The circles represent node names remaining in the temporary memory at these two stages in the production of a sentence. The numbers give the order in which rules have been applied

If, according to the grammar, indefinitely long sentences can be produced (or equivalently, if the set of sentences is infinite), how much temporary memory will have to be provided in the device? An unlimited amount? Let us examine this question. In Figure 4 on the left is shown a *regressive* structure. We call it a regressive structure because the machine has to go down the stem expanding 1, 2, 3, 4 and 5, storing a number of symbols in its memory (here, four, because there are four unexpanded branches), then it has to go back up, expanding in turn the branches growing from A, B, C and D. This regressive structure has a *depth* of four. The depth of a node is numerically equal to the number of symbols in tem-

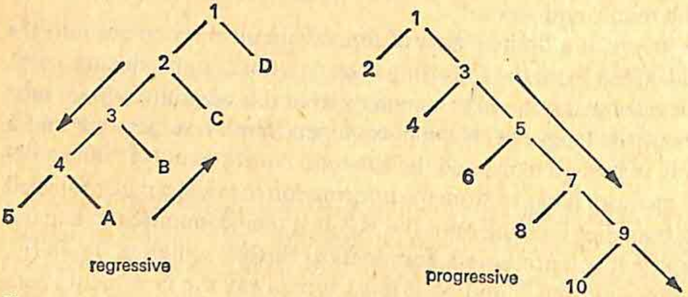


Figure 4 A regressive structure like the one on the left requires that an additional symbol be stored for each step down. This structure, which has a depth of four, requires that the four symbols A, B, C and D be stored. A progressive structure like the one on the right can be extended indefinitely by using a temporary memory that can hold one symbol only

porary memory when that node is about to be expanded. On the right is shown a *progressive* structure. The machine can continue down the main stem, expanding as it goes, never retracing its steps, and it puts only one symbol away in its temporary memory each time; and each time it takes it right out again and expands it. It is clear that as *regressive* structures grow longer they require more and more memory; but *progressive* structures do not. They can continue indefinitely with a minimum of memory.

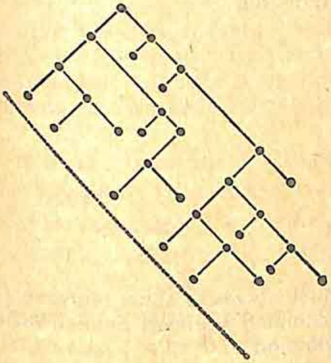


Figure 5 If the temporary memory can contain only three symbols, the structures it can produce are limited to a depth of three and can never penetrate the dotted line

There remains the picture illustrated in Figure 5. If the memory had room for only three symbols at one time, it could produce such a structure as is shown here, but it would not be able to produce any structure that would penetrate the dotted line. We have noted that a person can remember only about seven items and repeat them back. It turns out that the sentences of English can be represented by trees that are limited, as if by a dotted line, to a depth of about seven. It can thus be seen that the structure of English is inherently *unsymmetrical* in the sense that there does not exist the same possibility of branching off to the left as to the right. It is interesting that the notations of mathematics and symbolic logic do not have this restriction. They are essentially written languages. The problem of remembering future commitments is solved by the mathematician by looking back at what he has written; he does not have to keep it all in his mind as he goes along. But in a spoken language one has to keep track of things in one's head as one goes along, and one has the limitation, about seven items, on the span of immediate memory.

It is *not* the case that very deep structures are allowed in English but are just not used because of a memory limitation. It *is* the case that the grammar of English is so constructed that excessively deep constructions are actively prevented, and alternative constructions of lesser depth are provided. And it is this fact that accounts for much of the complexity of English over and above that of mathematical or symbolic logic notations.

One of the most striking features of English syntax is the coupling of a part-of-speech system with word-order rules. This is a case in point. Take, for example, the sentence 'Very clearly projected pictures appeared'. Here is a purely regressive structure with a depth of four.

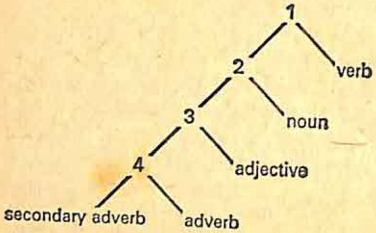


Figure 6 The part-of-speech hierarchy in English serves to count regressive steps and helps to prevent sentences from becoming too deep. English lacks such a counting system in the progressive direction

Figure 6 shows how the sentence is organized. Counting down from the top we have first a verb – 'appeared' – then a noun – 'pictures' – then an adjective – 'projected' – then an adverb – 'clearly' – and finally a secondary adverb of a special limited class – 'very'. It is difficult to build much further in a regressive direction. The part-of-speech system provides a method of automatically counting down a regressive branch and stopping the construction before it gets over the depth limit. It is clear that this system is a counting scheme and is not tied to the meaning of the words, because it is often possible to shift the words up and down the ladder by making appropriate changes in part of speech. We can have 'a very clear projection', where the participial adjective 'projected' becomes the noun 'projection' and the adverb 'clearly' becomes the adjective 'clear'. The thing that remains constant is the part-of-speech labeling that serves to keep track of the extent of regression. It is characteristic of English that it lacks such a means for counting in the progressive direction.

There are many progressive constructions that can run on without limit. First of all there is coordination, as in a long string of coordinated noun phrases. This can go on without limit in the progressive direction. Then there is the accumulative pattern of modification of several adjectives before a noun. But when we get to subordinate clauses we have to look at the matter in a bit more detail.

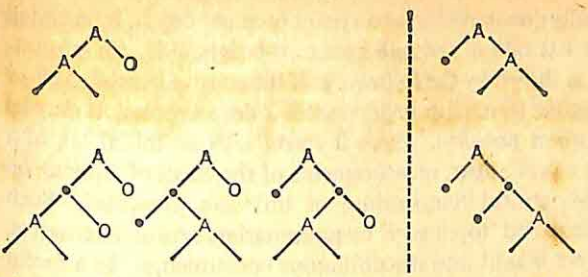


Figure 7 Indefinite reapplication of a rule A according to the patterns on the left would lead to structures exceeding any depth limit. They are ungrammatical. Indefinite reapplication according to the patterns on the right do not involve an unlimited increase in depth and are therefore safe

In Figure 7, various structures are shown that involve reapplication of a rule labeled A. Those at the left of the dotted line would be ungrammatical because they imply the possibility of unlimited application of the rule A, and would lead to an unlimited increase in depth. They would trap a person into going too deep. Those at the right could occur and would cause no trouble. In English there are a number of structures involving subordinate clauses that can be iterated indefinitely. They have structures like those on the right. For example, the object clauses: 'I watched him watch Mary watch the baby feed the kitten.' Or, another type of object clause: 'I imagined him hearing the announcer reporting Bill catching Tom stealing third base.' A 'that' clause: 'John said that Paul said that Bill said that he had won the race.' But a 'that' clause can also be used before the verb as a subject clause, and it is a regressive construction. It cannot be iterated indefinitely. 'That it is true is obvious.' Now consider two 'that' clauses: 'That that it is true is obvious isn't clear.' It is ungrammatical. It is the structure closest to the dotted line on the left in Figure 7. What is used, instead, is one of the typical complexities of English syntax, the anticipatory 'it': 'It isn't clear that it is obvious that it is true.'

And now here is another progressive structure that can be extended indefinitely, an object 'what' clause: 'He knows what should have been included in what came with what he ordered.' But if a 'what' clause is placed in the subject position, it can only be done once. Here is an attempt at using three 'what' clauses: 'What what what he wanted cost would buy in Germany was amazing.' It is ungrammatical. And it is a regressive structure. Instead, using three of the complications of grammar – a discontinuous constituent, a passive, and a nominalization – one can build a substitute progressive structure: 'It was amazing what could be bought in Germany for the cost of what he wanted.'

The discontinuous constituents in English seem to serve the purpose of postponing potentially deep structures to a point of lesser depth, by carrying structures from the left side of the tree over to the right side. An interesting example of this is shown by the following: If the passive is used to allow the subject of the active to remain unexpressed, a deep expression may be thrown into the subject position, where it starts with an initial tax of a depth of one. 'In a recent paper, measurements of the effect of alloying on the superconductive critical temperature of tin were presented.' Such structures have been called 'top heavy' by grammarians. As an alternative, frequently the subject is split into discontinuous constituents: 'In a recent paper, measurements (were presented) of the effect of alloying on the superconductive critical temperature of tin.' Here, by a complication of syntax, the prepositional phrase modifier of the subject is postponed until after the predicate, so that we have the grammatical but 'illogical' order of subject, predicate, modifiers of the subject. Time does not allow for more examples, but there are many.[2]

From the point of view advanced here, the sentences of English can be produced by a finite-state device. The syntax of English is not an endless catalog of whimsical complications, although it does contain some relics of the past. Neither does English appear to be an abstract formal system on a par with certain elegant mathematical notations. Instead, it is a particularly well engineered instrument of communication, with many ingenious innovations for adapting it to the capabilities of its users and to circumvent as much as possible the limitations of the human mind.

It remains to be seen how well the hypothesis applies to other languages. I hope that some of you will become interested in making such tests. Of course other languages are not the same as English. They will not have the same features as those that we can explain in English on the basis of the depth hypothesis. They may have just the opposite features to some of those explained in English – postpositions instead of prepositions, for example. But it is the over-all structure of the syntax of the language that should be judged in the light of the hypothesis, since it is usually the cooperation of a number of features that keeps any given sentence from becoming too deep. It is to be expected that all languages have methods of restricting regression, methods of conserving depth, and methods of circumventing the memory restriction so as to maintain expressive power.

The last part of the hypothesis proposes that depth phenomena be added to the list of already known factors that affect language change. A depth factor in language change should be easily observable if it exists, and it will probably show up quickly when grammatical sketches of different periods in the development of a language are compared.

---

[2] See footnote 1, p. 115.

References

MILLER, G. A. (1956), 'Human memory and the storage of information', *IRE Trans. Information Theory*, vol. IT-2, no. 3, pp. 129–37.

YNGVE, V. H. (1960), 'A model and an hypothesis for language structures', *Proc. Amer. philos. Soc.*, vol. 104, pp. 444–66.

9 C. Watkins

Preliminaries to the Reconstruction of Indo-European Sentence Structure

C. Watkins, 'Preliminaries to the reconstruction of Indo-European sentence structure', *Proceedings of the Ninth International Congress (1962) of Linguists*, 1964, pp. 1035–42.

Syntax is notoriously ignored in most studies of comparative and historical linguistics. When one turns from phonology and morphology to syntax, what is conspicuously lacking in any of the manuals are syntactic equations, of the sort on which one can build reconstructions. Yet there is only an apparent, not a real contrast between e.g. $*p > f$, and a comparable putative syntactic historical filiation. It would be absurd to think that the physical body of a phrase or similar utterance can be historically transmitted. But the underlying structure of a phrase, as in our case the rules of arrangement of the elements, can be presumed to constitute a linguistic system functioning in time, and as such to be susceptible to analysis by the comparative method as well as by other techniques of historical linguistics. The 'tractability' of the syntactic system for historical investigation is only in degree different from that of the phonological or the morphological.

In this paper we seek to establish certain equations in phrase structure among certain of the older IE languages: Vedic, Hittite, Latin, Greek and Old Irish. The elements of the sentence or phrase with which we will be concerned are four: the sentence connective (N), the enclitic pronominal element (E), the preverb (P) and the finite verb form (V). Some or all of these four elements may be combined to form what we may term the 'verb phrase' of IE, and of the later historically attested languages, in independent or 'principal' clauses. Our concern here may be described by the title of a significant article by Dell H. Hymes, 'Positional analysis of categories: a frame for reconstruction' (1955). Put more loosely, it is the problem of 'word order', though it is clear in IE, and even clearer in Athapaskan, that the elements in question are not all 'words'. It is not of particular import whether we speak of this as a problem of syntax, or as a problem of morphology, as does Hymes; I prefer to call it syntax, since it is traditionally put there in IE comparative grammar.

It will be evident, I hope, that the formulae given below for the relative positions of different elements in the sentence in the various languages is not meant to constitute a set of hard and fast rules for these languages, to which there are next to no exceptions. Such an assumption in the field of syntax would be of course hopelessly unrealistic. But it is nonetheless an observable fact that, as J. Gonda has put it, 'The grammatical procedures by which words and word groups were arranged and united into meaningful larger units were often highly conventional' (1959, p. 8). We will be concerned with such conventional patterns of word order in the languages under consideration.

These recurrent patterns may and do vary from language to language. Yet such patterns may also exhibit, among a set of languages, striking recurrent features of similarity. Where we know from other considerations that these languages are cognate, though this condition is by no means a necessary one, it is a legitimate hypothesis that such recurrent similarities of syntactic patterning may result from genetic filiation. This is of course the basic assumption behind any attempt at historical and comparative syntax, and has been with us since the early days of IE studies. Yet the atomistic approach of the neo-grammarians left its imprint on the study of word order as well. For Delbruck, the position of V was a separate question from that of the position of P, and of E. To state that the normal position of the verb in IE was sentence-final, as he does, hardly serves to characterize the IE phrase; sentence-final position for the finite verb is equally normal in, for example, Burushaski.

To the earlier atomistic approach to linguistic history we have for some time opposed the notion of language as a system of interrelated units, on whatever level; and this view must be applied to word order just as to phonology. In our concern here for the four elements of the verb phrase mentioned above, what is of significance is not the position of each element, but the position of each relative to the others, and the ensuing sentence patterns which we can formulate. It is on this basis that one can, I think, make a meaningful statement about IE phrase structure. So, for a simple case, we may contrast the IE type #.E...V#, with verb at the end and pronominal object in second position, to the Burushaski type #...EV#, considering the Burushaski prefixal pronouns (obj. of trans. verbs, subj. of intrans. verbs), as E. The picture is entirely different from that to be gained from examining the position of V alone.

To take another example: one of the few generally accepted syntactic statements about IE is Wackernagel's Law, that enclitics originally occupied the second position in the sentence. Wackernagel based his formulation on the evidence of Greek, Indo-Iranian, and Latin, and recognized its application to Celtic (see p. 131); later investigations of Balto-Slavic, and in particular the evidence of the subsequently discovered Anatolian languages, have abundantly confirmed his view. Kuryłowicz, however, has suggested (1958, p. 613) against V. V. Ivanov (*l.c.*), that Wackernagel's Law belongs on the plane of a general linguistic feature rather than to the comparative grammar of the IE languages alone. I would disagree with this statement in any case; but what is decisive is not the fact that an enclitic (particle or pronoun) in the earlier IE languages usually occupies second position. This might conceivably result fortuitously from general conditions of phrase contour. But when we observe the same fixed order among two or more such enclitics in all the older languages; when unaccented connectives such as *k^we, *de (Gk. δὲ, OIr -d-), *yo (Hitt. -ya, OIr. -e, -a-) always precede and never follow unaccented pronouns in Indic, Greek, Latin, Hittite with the other Anatolian languages, and Old Irish; then we are forced to assume that there is in fact a direct historical relation. The factor of chance can be ruled out.

For the situation in Vedic, we may reproduce the remarks of Gonda (1959, p. 7):

> In the majority of the works examined, the final position [of the verb] is most usual. Although the initial position has never been predominant, the verb may, speaking generally, occur at the beginning in a well-definable number of cases.

It is the great merit of Gonda's work to show that where the verb preceded by the object or subject is not in final position, such sentences are usually 'amplified', i.e., 'they are from their beginning until the verb complete in themselves'; the remainder may be left out without affecting the structure of the sentence. This provides additional corroboration for the basic final position of V; it is a pattern which recurs in other IE languages as well. The position of P has been described independently by Bonfante (1930) and Kuryłowicz (1936): either it occupies initial position, separated from V ('tmesis') or else immediately precedes V, still written as a separate word. In subordinate clause we have already in the Rig Veda the 'univerbation' of juxtaposed P and V, which becomes the rule later for principal clause as well, with the elimination of tmesis. This indicates that P and V even in tmesis are constituents of a single semantic 'word'. In both cases V occupies final position, with the reserve for the 'amplified sentence'. Enclitic particles and pronouns (E), when present, occupy regularly the second position, as established by Wackernagel. Hence we may set up the following formulae for the Vedic sentence or clause, marking the boundaries by #.

Simple verb:

(a) #V(E)...#(marked);
(b) #.(E)...V#(unmarked);

Compound verb:

(c) #P(E)...V#;
(d) #.(E)...PV#.

We shall see that these patterns repeat themselves with extraordinary regularity in the other IE languages under consideration. Rigvedic examples:

(a) *bháranti vām. . .*(1.151.8);
(b) *.īm. . .jabhāra* (4.27.4);
(c) *prá vām. . .bharante* (7.72.4);
(d) *.te. . .prá bharāmasi* (8.66.11).

There are also cases of V (which is unaccented in principal clauses) occupying second, i.e. enclitic, position. It may be preceded by P in initial position, and most frequently by the connective pronoun *sá* (N). When an enclitic pronoun (E) is also present, this precedes enclitic V (though the situation is often not distinguishable from an amplified sentence). For example, *antáḥ paśyanti vṛjinā́* 'they see into evil' (2.27.3, #PV. . .#); *sá veda yajñám* '(and) he knows the sacrifice' (3.11.1, #NV. . .#); *etā́ vo vaśmi údyata* 'I wish you these proffered (words)' (2.31.7, #.EV. . .#).

As Delbrück has noted (1893, p. 51), where both enclitic particle and enclitic pronoun are present in second position, the particle precedes the pronoun: *dyáuś ca tvā. . .*(3.6.3). Here *ca* > $*k^{w}e$ is properly an enclitic connective N. The pattern (NE) in enclitic position is likewise general in the other languages. Connective N may also occur accented; in this case N occupies initial position. The clearest evidence for initial N is furnished by Hittite, for which see below; but in the frequent concatenation of phrases by the pronoun *sá-/tá-* in initial position, Vedic has preserved a clear reflex of the older situation. The pronoun *sá* as N may be followed by E (*sá u* passim); including enclitic V (supra); more usually V is final. When V is a compound, the formula is always #N(E). . .PV#: *sá devó devā́n práti paprathe* '(and) the god has expanded as far as the (other) gods' (2.24.11). The older invariant form *sá*, meaning simply 'and', is preserved in the conjunctions *sá yádi, sá yátra, sá céd* (Wackernagel, 1955, pp. 257–61). The reconstructible **so* was termed by Wackernagel a 'blosses Fulcrum'; it is simply a fossilized connective N.

The order of elements in the Hittite sentence is in general far more rigid than that of Vedic. The normal position of V is sentence-final, but initial position exists here also as a stylistically marked variant. Both single E and 'strings' of E (pronouns and particles) always occupy the second position, and are written as one word with the initial element of the sentence, whether this be N, P, V, or other lexeme. Hittite knows semantic composition of P and V, though the actual accentual 'univerbation' has not yet taken place (as it has in the more recent Lydian, for example), and the two are always written as separate words. The far commoner position of P is immediately before V in final position; but examples are not rare of P in sentence-initial position, with V at the end. For sentences with P, E, V, then, we have the same four formulae as in Vedic:

(a) #V(E). . .#;
(b) #.(E). . .V#;
(c) #P(E). . .V#;
(d) #.(E). . .PV#.

Examples from the Laws:

(a) *kuenzi-ma-an* LUGAL-*uš* 'but the king can kill him'(II § 74);
(b) *tepu-ši išḫiyanzi* 'they bind little on him' (I § 94);
(c) *šēr-wa-ši šarnikmi* 'I shall make restitution for him' (I § 95);
(d) *išḫas-šeš-a šēr šarnikzi* 'and his master makes restitution' (I § 99).

Among the enclitics in second position there is a fixed order of precedence; cf. Laroche (1958) for Hittite, Luvian and Hierogl. 'Hit.'. We may note here that the enclitic connective -*(y)a*, the functional equivalent of IE *$k^{w}e$ (cf. *kuiš-a* = Lat. *quis-que*), always precedes enclitic pronoun: KUR UGU-*ya-mu*. . .(Ḫatt. I 26), with the same order of enclitics (NE) as in Vedic.

It is for the sentence connective N in initial position that the Hittite evidence is uniquely valuable. In archaic Hittite, we have three connective particles, *nu*, *ta*, *šu*, which always occupy initial position, and serve to link the sentence or clause with what precedes; they may be translated 'and', or simply ignored in translation. Though the later language uses the connective (*nu* alone) with greater regularity than the earlier language, it is clear from our earliest texts that N in initial position, optionally followed by one or more E, is a solidly established syntactic feature. As long recognized, these forms represent deictic pronominal stems; but they show the bare stem alone, IE **nu* **to* **so* **e* (> Luvian *a-* = Hitt. *nu*), and are formally on an older level than the corresponding inflected pronouns themselves. Formally, *nu* has been equated with the Old Irish 'empty' preverb *no*; we can make a syntactic equation as well (see Dillon, 1947). Both occupy sentence-initial position and both are regularly followed by E. In both Hittite and (archaic) Old Irish, V is final; we can juxtapose Hitt. *nu-mu* ^{d}*IŠTAR. . .kaniššan ḫarta* 'and Ishtar held me in favor' (Ḫatt. I 66) and OIr. *no-m Choimmdiu -coima* 'the Lord cherishes me' (*Early Ir. Lyr.*, 2.2). The formula for both is #NE. . .V#. Where V is a compound, it is preceded in final position by P; the formula is thus #N(E). . .PV#. just as above in Vedic. For example, *nu-kán tamedani kuedanikki andan paitteni* 'you go in to someone else' (see Schuler, 1957, p. 19), which may be contrasted with *anda-kán tamendani lē kuedanikki paitteni* 'do not go in to someone else' (Schuler, 1957, p. 25), showing #PE. . .V#, without N. As we shall see below, there is ample evidence in the other languages to justify attributing this type of construction with N to IE itself.

We may consider now the situation in Latin. Here the final position of V has been long recognized as basic, and we may state in general that the older the text, the more regular this appears; for example, the SC de Bac., or the Columna Rostrata, where it is constant. On the other hand, there is good evidence for permissible initial position of V as a marked variant (Kroll, 1918). For the original second position of E (though this 'law' ceased to operate during the course of the historical period) it is sufficient to refer to Wackernagel's study. I point out only that earlier Latin as well regularly presents the order *-que*+pronoun (NE) among the enclitics, e.g. *dantque eum* (Enn. Euhem.). By the time of our earliest extended records, the univerbation of P and V is an accomplished fact, and initial or final position of V includes V > PV. But clear traces remain of an earlier stage preceding the univerbation; that commonly described as 'tmesis'. For our type (d) #.(E). . .PV#, for example, *manum endo iacito* (XII Tab.); for type (c) #P(E). . .V#, *prae tet tremonti* (Carm. Sal.), *sub uos placo* (Fest.). The latter are minimal verb phrases, in that nothing comes between E in second position and V in final position, and the pattern PEV comes to be normal for literary tmesis; but I suggest that we have the real separation in *ē nōs Lasēs iuuāte* (Carm. Arv.), where I take *ē* as the preverb (: Skt. *ā́*) appearing also in Lat. *hērēs* > **ghēro-+ē-d-*: Skt. *ā-dā-*.

One of the striking characteristics of the earliest Latin prose is the almost universal use of *-que*, enclitic on the first word of the sentence or clause, which functions as a pure sentence connective, and not as the copulative conjunction of the classical language. Compare the inscription of the Columna Rostrata, of the Carmen Euocationis, where enclitic N *-que* is as constant as sentence initial N *nu* in classical Hittite. This stylistic feature shows the antiquity in Latin of sentences involving N. We have however in Latin one extraordinary archaism which proves the former existence of an initial N in Latin, of the same form as Hitt. *nu*: the adverb *nudius tertius* 'day before yesterday'. Hofmann has shown that this is a nominal sentence, '(now) it is the third day' (1924, p. 77). We can set it beside Hitt. *nu* ITU 10 KAJ (Otten, Kummarbit 7) 'and it is the tenth month', which is an identically constructed sentence; the equation shows that Latin once possessed an N identical in form, function, and position with Hitt. *nu*.

For Greek we are faced at first sight with a bewildering variety of positions; there is little to be made from such contrasts as Hecataeus 1 F 1 τούτου' δ' Οἰνεὺς ἐγένεο: Οἰνέως δ'ἐγένετο Αἰτωλός. See in general Dover (1960); Greek would appear to have gone farther than any other IE language in the elaboration of a 'free' word order. For V, however, the basic constituents would appear to be final position, initial position, and second (enclitic) position. The last is particularly notable in Mycenean, after N (*ho*, see below) in initial position: *ho-deksato*. . ., #NV. . .#, which agrees with this type in Vedic.

The univerbation of P and V is older than Mycenean times (*apudōke* etc.), but the separation of the two was preserved in the literary language. Wackernagel points out that 'Von alters her tritt Tmesis am ehesten ein, wenn das Präverbium zugleich an der Spitze des Satzes steht' (1926, p. 174), which is precisely our type (c) #P(E). . .V(. . .)#, e.g. κατὰ πίονα μηρί' ἔκηα (A 40). The other Homeric type of tmesis, where P immediately follows V, as in τότε δ' ἤδη ἔχεν κάτα γαῖα μέλαινα (B 699), I suspect to be a (purely literary?) Greek innovation, based on the inherited syntagm noun + postposition (see Hittite). It would have replaced our type (d) #.(E). . . PV(. . .)#, with P immediately before V. The occasional presence of a caesura noted by Sommer between apparently 'univerbated' P and V (e.g. ἀμφι|πονησόμεθ' Ψ 159) proves the original existence of a word boundary between P and V in type (d) in Greek, exactly as in Vedic and Hittite. The same is shown by the presence of a word divider between P and V (with V in final position) in Cypr. εξ|ορυξε (Schwyzer, 1939, p. 679).

The second position of E in early Greek has been amply documented by Wackernagel, and confirmed by Dover. Delbrück has furthermore noted, after Monro, that enclitic connective (e.g. δὲ) in second position precedes enclitic pronoun (NE); the same is characteristic of Mycenean, e.g. *dāmos-de-min phāsi*, *ekhei-de-min* (Ventris and Chadwick, 1956, p. 254) (cf. *dantque eum*). The former example appears to show an enclitic V following E, an enclitic series (NEV) beside the latter V(NE); it may be compared with Vedic (EV).

It is Mycenean which preserves connective *ho* in its original form, function, and syntactic pattern: the particle *o-*, with a graphic variant *jo-*, which regularly begins a clause, and is always written together with the following word, usually a finite verb: *ho-wide* 'and he saw', *ho-ophēlonsi* 'and they owe' *ho-agrēse* 'and he took' (Ventris and Chadwick, 1956, p. 43). The negation of the simple verb *didonsi* is *ou-didonsi*; but the negation of *ho-agrēse* is *ou-*k^w *e agrēse*, which proves that *ho-* and k^w*e* have the same semantic content, that of a connective 'and'. The same conclusion is offered by comparison of sentences of the type NN *mo(i)roppās toto wetos ho-agrēse* ZE 1, with those of the type NN *theoio doelā ekhei-*k^w*e onāton*, which are of identical structure: 'NN (is) shareholder this year; and he took one pair'; 'NN (is) the servant of the god; and she holds a lease'. Even the convention of writing *ho-* as one word with what follows is exactly paralleled in Old Hittite usage with the connective *nu-* (Otten, 1953, pp. 59–61). The great stylistic development of the use of connective particles in classical Greek thus had its origins in IE times; it may indeed be compared with the much humbler generalization of *nu* in classical Hittite.

We may take Old Irish as 'tertium' comparationis; it is of critical importance for its isolated position on the western fringe of the IE area, and its evidence for comparative syntax is essentially untouched. In classical Old Irish, V occupies sentence initial position; with compound verb, the complex P+V occupies the same position; pronominal objects (E) are always unstressed, and may occur either 'infixed' or 'suffixed'; suffixed E occurs after uncompounded V, and infixed E occurs between P and V. The respective formulae are thus #V(E). . .# and #P(E)V. . .#; it is evident that the position of E is consistently second in the sentence, and that we have a clear conservation of Wackernagel's Law. The same fixed order of enclitics (NE) recurs in Old Irish as well, e.g. _nachim-_ < _*ne k_^{w}_e me_, or the class B and C pronouns, e.g. _-dam-_ < _*de me_ (_-d_: δὲ). For OIr. _no_, the lexically empty preverb serving as 'prop' to infix pronouns to a simple verb, which continues the IE connective _*nu_ in initial position, see the section on Hittite above (p. 128). Not only have we _no_: Hitt. _nu_ continuing N, but also _se-_: OHitt. _šu_ as well. Cf. _se-ch is_ = _no-ch is_; the apophony _*se/so_ is that of Lat. _olle_ < _*ol-se_: Gk. ὁ < _*so_.

Now besides the above patterns for P and V in classical Old Irish, archaic Old Irish preserves two earlier syntactic patterns, which have so far gone unnoticed in comparative IE studies; yet together they exactly recover the situation in the other early IE languages. These are Bergin's Law (Bergin, 1938), and 'tmesis'. In tmesis, P(E) occupies initial position, and V is sentence-final; in Bergin's Law, both simple V and compound PV are sentence-final. It is clear from the comparative data that initial position of V is inherited; we can combine this with tmesis and Bergin's Law, and the result is four sentence patterns for the earliest Old Irish:

(a) #V(E). . .#;
(b) #. . .V#(B's Law);
(c) #P(E). . .V#(tmesis);
(d) #. . .PV#(B's Law).

One can finally show that the classical Old Irish type #P(E)V. . .# is a late development, representing merely the univerbation of the tmesis type (c).

We have considered the relative positions of four elements of the verb phrase in five earlier IE languages. The situation is such that there can be no question of the patterns of order of any one of these languages influencing the pattern of any of the others, in respect of the elements considered. For this reason the essential identity among the five languages in the patterns of structure of the verb phrase is sufficiently remarkable that we may safely exclude chance; we have to deal here with direct syntactic inheritances from common IE times. I would submit, then, on the evidence of the equations among all these five languages for the earliest discernible patterns of phrase structure, that we are entitled to reconstruct the following types for the IE verb phrase:

1. For the uncompounded finite verb, an opposition of sentence-initial and sentence-final position of V, with a permissible E or series of E in second position: (a) #V(E). . .# (b) #.(E). . .V#. The latter was evidently 'normal', and the former the stylistically marked member of the opposition.

2. For the finite verb semantically compounded with a preverb, obligatory final position of V, and an opposition of initial and prefinal position of P, again with permissible E or series of E in second position: (c) #P(E). . . V# (d) #.(E). . .PV#.

3. For sentences in narrative or continuous style, a connective particle may appear in initial position, followed by a permissible E or series of E in second position, with V in final position, preceded by P in the case of a compound verb: (e) #N(E). . .(P)V#.

4. In all five types (a–e), a connective particle (N) may also appear in second position; in this case an enclitic pronoun (E), if present, always comes after enclitic N. Thus we have permissible phrases beginning #.N(E), #PN(E), #VN(E), and even #NN(E), as in **so/e k*^{w}*e* (Ved. *sá ca* = OIr. *se-ch*), **nu k*^{w}*e* (OIr. *no-ch* = Goth. *nauh*).

For the enclitic (second) position of V, we have the evidence of Vedic and Greek. But since these are precisely the only two IE languages for which we have other good evidence for the atonic character of V in non-initial position, it is possible that this position is an independent innovation in both; see Kuryłowicz (1952, para. 4).

It is obvious that we have not considered the evidence of all the IE languages, nor all the evidence in those we have treated. Iranian agrees essentially with Indic, and both Baltic and Slavic would afford striking confirmation. The historical transformations in Germanic require particular investigation. I have purposefully ignored the situation of the dependent or subordinate clause, for reasons of space; but in this syntactic type the same equations can be made among the same languages, with the addition of the 'relativing' element **yo-*, **k*^{w}*o*. This element is ultimately only a member of the class N, appearing in either initial or second (enclitic) position. Likewise worthy of investigation is the treatment of the negation, which closely parallels P in distribution. Wackernagel (1928, p. 172) notes as 'eigentümlich' Goth. *miþ-ni-qam* 'did not go in with' 'οὐ συνεισῆλθεν' (John VI 22); but note the normal Hitt. type *anda UL pait* 'did not go into'.

It is worthwhile noting that just as we have tonic and atonic personal pronouns (our E concerns only the latter), the connective N may appear in either initial or second (enclitic) position. So **nu* is initial in Hitt. *nu*, OIr. *no*, Ved. *nú*, but enclitic in Arc. ὅ-νυ, Ved. *nú*; **to* is initial in Hitt. *ta*, Goth. *þauh* > **tu-k^{w}e*, enclitic in Hitt. *na-tta*, Ved. *tú*; **so* is initial in Hitt. *šu*, Gk. *ho-*, enclitic in Hitt. *na-ššu*. Similarly Lat. *quī* but *-que*, Gk. ὅς but Hitt. *-ya*, OIr. *-e*.

We may add finally that the doctrine of the sentence connective permits the explanation of a peculiar IE dialect feature: the augment *e-*. We have seen that OIr. *no* corresponds formally and functionally to the Hittite connective *nu*. But there is a secondary function of *no* in Old Irish: it always (except in archaic poetry) precedes verb forms which have (OIr.) secondary endings. In this function it is identical with the augment *e-* of Greek and Indo-Iranian. Luvian provides the missing link: Luvian *a-* corresponds functionally to the Hittite connective *nu*, and formally to IE *e-* (cf. Luv. *as-* 'be', *ad-/az-* 'eat'). We may set up a square as follows: illustrating both the formal and functional correspondences:

OIr. *no* ———— Hitt. *nu*
IE **e* ———— Luv. *a*

The Luvian connective *a-* may thus be exactly equated with the 'IE augment', which is thus revealed to be by origin a connective used with 'injunctive' forms in the function of continuous narration.

References

Bergin, O. (1938), 'On the syntax of the verb in Old Irish', *Eriu*, vol. 12, pp. 197–214.

Bonfante, G. (1930), 'Proposizione principale e proposizione dependente in indoeuropeo', *AGI*, vol. 24, 2, pp. 1–64.

Brugmann, K., and B. Delbrück (1893–1900), *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*, Strassburg, K. J. Trübner.

Delbrück, B. (1893), *Vergleichende Syntax der indogermanischen Sprachen*, vol. 3, Trübner.

Dillon, M. (1947), 'Celtic and the other Indo-European languages', *TPhS*, pp. 15–24.

Dover, K. J. (1960), *Greek Word Order*, Cambridge University Press.

Gonda, J. (1959), *Four Studies in the Language of the Veda*, Mouton.

Hofmann, J. B. (1924), 'Syntaktische Gliederungsverschiebungen im Latein infolge Erstarrung ursprünglich appositioneller Verhältnisse', *IF*, vol. 42, pp. 75–87.

Hymes, D. (1955), 'Positional analysis of categories: a frame for reconstruction', *Word*, vol. 11, pp. 10–23.

Ivanov, V. V. (1958), 'The importance of new data concerning Hittite and Tocharian languages, etc.', *Proc. Eighth Internat. Congress Linguists*, Oslo University Press, p. 611–13.

Kroll, W. (1918), 'Anfangsstellung des Verbums im Lateinischen', *Glotta*, vol. 9, pp. 112–23.

KURYLOWICZ, J. (1936), 'Les formes verbales composées du Rigvéda', *Bulletin de la Société de linguistique Polonaise*, vol. 5, pp. 39–46.

KURYLOWICZ, J. (1952), *L'accentuation des langues indo-européennes*, Krakow, Nakl. Polskiej Akademii Umiejetności.

KURYLOWICZ, J. (1958), 'Discussion of V. Ivanov's report', *Proc. Eighth Internat. Congress Linguists*, Oslo University Press, p. 613.

LAROCHE, E. (1958), 'Comparaison du louvite et du lycien', *BSL*, vol. 53 pp. 159–97.

OTTEN, H. (1950), *Mythen vom Gotte Kumarbi, Neue Fragmente*, Berlin, Akademie-Verlag.

OTTEN, H. (1953), 'Die inschriftliche Funde', *Mittheilungen der Deutschen Orient-Gesellschaft*, vol. 86, pp. 59–64.

SCHULER, E. von (1957), *Hethitische Dienstanweisungen für höhere Hof- und Staatsbeamte*, Graz, Ernst Weidner (Archiv für Orientforschung Beiheft, vol. 10).

SCHWYZER, E (1939), *Griechische Grammatik*, Munich, Beck.

VENTRIS, M., and J. CHADWICK (1956), *Documents in Mycenaean Greek*, Cambridge University Press.

WACKERNAGEL, J. (1926–8), *Vorlesungen über Syntax*, Basel, E. Birkhouser.

WACKERNAGEL, J. (1955), *Kleine Schriften*, Göttingen, Vandenhoeck u. Ruprecht.

10 M. Ivić

Non-omissible Determiners in Slavic Languages

M. Ivić, 'Non-omissible determiners in slavic languages', *Proceedings of the Ninth International Congress (1962) of Linguists*, 1964, pp. 476–9.

1. A determiner (D) of a given grammatical unit (X) is to be called omissible if the phrase D+X may be transformed into Ø+X, without changing the grammatical value of X. For instance, the Serbocroatian Accusative form *devojku* in *vidim lepu devojku* 'I see the beautiful girl' remains in the same grammatical function of direct object even if the phrase *lepu devojku* becomes transformed into Ø+X: *vidim devojku* 'I see the girl'.

Thus the determiner is of the OMISSIBLE type if:

$$D+X \rightarrow \emptyset+X = X$$

2. NON-OMISSIBLE determiners are qualifiers and quantifiers which cannot be omitted without specific grammatical consequences. Either (instance 1) the transformation of D+X into Ø+X cannot take place for it would lead to nonsense or (instance 2) the transformation of D+X into Ø+X would yield a new grammatical unit (Y).

Instance 1: The Czech example with genitivus qualitatis *dívka černych vlasů* 'the girl with black hair' cannot undergo the transformation test, the phrase **dívka vlasů* (=Ø+X) being ungrammatical.

Instance 2: The Serbocroatian phrase *celo proleće – čekao je celo proleće* 'he was waiting the whole spring' may be transformed into Ø+X, but that would mean the transformation of the temporal quantifier into the direct object: *čekao me proleće* 'he waited for spring'.

3. In instances concerning the use of the omissible determiner the relationship of D with X may be described as a simple conjunction of two grammatical units in the speech string (D+X = D+X). As far as the phenomenon of non-omissible determiners is concerned, the conjunction of the two grammatical units D and X is used as a specific morphological device for the identification of a third grammatical unit (Z): D+X = Z.

Thus:

Phrases with non-omissible determiners function as minimal grammatical units and that fact should be correctly observed and described in Slavic grammars.

The grammatical category of non-omissible determiners is to be found, of course, in non-Slavic languages too. But the Slavic languages are characterized by a comparatively wide use of non-omissible determiners in various situations.

4. It is noteworthy that in many instances in which grammars of Slavic languages mention the archaic free use of cases, we are concerned with a misunderstanding based on the fact that the existence of the grammatical category of non-omissible determiners has not been recognized. Examples quoted in this respect concern usually the case-form without prepositions which belongs nevertheless to the bound type being obligatorily used with a determiner. We should remember, for instance, the archaic 'free' (=without prepositions) use of Instrumental forms denoting quality in examples like Old Russian: *bě že Mьstislavъ debelъ tělomъ, čermenъ licemъ, velikyma očima* 'Mstislav was thickset and ruddy-faced, with big eyes'.[1] The phrase *M. velikyma očima* could not be transformed into **M. očima*. The same rule of transformation is valid even today, as may be shown by examples from certain Serbocroatian dialects in which the old case-category of this type is still in use (see Ivić, 1954, pp. 216–21).

The Genitive denoting quality has occurred exclusively with non-omissible determiners in Slavic languages, from the very beginning of the historical period till now.[2] (Serbocroatian: *devojka crnih očiju* 'the girl with black eyes' → **devojka očiju*). The Genitive denoting possession appeared from the very beginning side by side with possessive adjectives. As is well known, in Old Slavic the repartition of their use was established following the principle: the Genitive form appeared only if the determiner was present with the noun, the adjective if the noun was used alone, without determiners (so *synъ boga žifago* 'the son of living God' as opposed to *synъ božii* 'the son of God') (Trubetskoy, 1937, p. 16). The same principle is more or less respected in some modern Slavic languages (especially in Serbocroatian)[3]. The genetivus temporis, used without prepositions, requires also the presence of a non-omissible determiner (Serbocroatian: *udala se prošle godine* 'she got married last year', but never **udala se godine*). The same is valid for the Accusative denoting time or quantity (if used without prepositions): Serbocroatian *spavao je ono veče* 'he slept that evening' but never **spavao je veče*; *radio je celu nedelju* 'he worked the whole week' but never **radio je nedelju*. And so on.[4]

---

[1] For more examples see Potebnja (1888, p. 147).

[2] There are only a few examples of exceptions to this to be found in Polish; see Mirowicz (1949, p. 77).

[3] The question whether a bound Genitive form is used as qualifier or to denote possession can be resolved by transformation test. In both cases the phrase with the Genitive form becomes transformed by means of a grammatical device which helps to convey the meaning 'being in possession of'. If the substantive with the Genitive mark becomes recognized as the possessor, we have to do with genetivus possessivus. If it is not the substantive in the Genitive form but the head of the phrase which is recognized as the possessor by the transformation test we have to do with genitivus qualitatis. Cf. the transformation of the following Serbocroatian phrases: (I) *crna kosa moje sestre* 'my sister's black hair' → moja sestra ima crnu kosu (=genetivus possessivus)/ (II) *moja sestra crne kose* 'my sister with black hair' → moja sestra ima crnu kosu (=genetivus qualitatis).

5. The grammatical relevance of the conjunction labelled $D+X$ has to be observed also within the scope of verbal categories.

The Serbocroatian phrase *dodje ponekad* 'he comes sometimes' with the PERFECTIVE present tense *dodje* (where *dodje* $= X$, and *ponekad* $= D$) cannot undergo the transformation into ${}^*\varnothing+X$ for it would lead to grammatical nonsense (*dodje* being not recognized as a minimal unit). On the contrary, the phrase *dolazi ponekad* 'he comes sometimes', with the IMPERFECTIVE present tense form, may undergo the transformation test which would yield a change of the grammatical meaning (following the pattern: $D+X \rightarrow \varnothing+X = Y$): *dolazi ponekad* 'he comes sometimes' ~ *dolazi* 'he is coming'. As may be shown by transformation tests, in the majority of South Slavic languages the present tense forms of perfective verbs do not occur except in phrases. Thus, for such languages, we must correct the well-known statement (usually quoted in Slavic grammars) that every Slavic finite verbal form automatically may function as a sentence. We have to add: but as to present tense forms of perfective verbs, they constitute a sentence *only if* occurring in phrases (for otherwise they cannot be taken as minimal units, either on the morphological, or on the syntactic level).

Just one more example as illustration (this does not exhaust the actual list of instances when Slavic verbs are used with non-omissible determiners): the Serbocroatian past gerund forms of PERFECTIVE verbs may be used free (= without determiners) but the past gerund forms of IMPERFECTIVE verbs may occur only in phrases, being obligatorily bound to the presence of a quantifier (*čekavši*/impf. verb/ *dva sata, otišao je kući* 'having waited for two hours, he went home', **čekavši, otišao je kući*).[5]

6. All the different examples already mentioned belong either to the type $D+X \rightarrow {}^*\varnothing+X$, or to the type $D+X \rightarrow \varnothing+X = Y$. It would be a contribution to the structural characteristics of the language under investigation to state all instances in which an opposition between bound and free forms may or may not occur. Even more, elaboration of any similar kind of grammatical phenomena, which have not hitherto been correctly observed,

---

[4] For more examples (from different Slavic languages) see Ivić (1959).

[5] More examples are quoted in Ivić (1962).

would certainly provide new information about the intersection of the morphological and syntactic levels.

7. The intersection of the grammatical and lexical levels would also be seen under a new light by pointing out correctly the grammatical phenomenon concerning the obligatory use of a determiner. As a matter of fact, it is not only the presence of the determiner which may be a relevant grammatical condition. It is even the choice of the lexeme used as non-omissible determiner which may have to be grammatically determined. In the example *čekao je celo proleće* the Accusative denotes the quantity of time. But if instead of the quantifier *celo* we put another adjective, like *sunčano* 'sunny' or *lepo* 'beautiful', the object-meaning of the free case-form remains unchanged: *čekao je proleće*, or *sunčano proleće*, or *lepo proleće* = Acc. of object ∼ *čekao je celo proleće* = Acc. of temporal quantifying.

The grammatical criteria by which a similar sort of lexical choice happens to be determined may be changed in the course of time (in Serbocroatian, for instance, one century ago it was possible to use examples like *desilo se nemačkog rata* 'it took place during the war with the Germans', now only: . . . *ovog rata* '. . . during this war', . . . *prošlog rata* '. . . during the last war' and the like, but never *. . . *nemačkog rata*). Slavic diachronic grammar should throw more light on such phenomena.

8. General linguistic theory may be interested in the fact that it is only the marginal use of a flexional form which is usually marked by the presence of a non-omissible determiner. We can mention here the use of Slavic case-forms as an illustration. There are three basic morphological possibilities to indicate case meanings in Slavic languages. They are: the free case-form, the case-form accompanied by a preposition, the case-form obligatorily accompanied by a determiner (see Ivić, 1961). The case-forms which by no means can have adjectives or adverbs as their substitutes must be used free. Adjectives or adverbs may be used as substitutes only in some instances but not invariably if the case-form is bound to the presence of a preposition. As to the case-forms requiring a non-omissible determiner, they may not only undergo in some instances the same kind of substitution, but they even must, without exception, belong to the same distributional class as qualifiers and quantifiers.

References

Ivić, M. (1954), *Značenja srpskohroatskog instrumentala i njihov razvoj*, Beograd, pp. 216–21.

Ivić, M. (1959), 'Srpskohrvatski padežni oblici obavezno praćeni odredbom kao promócnim morfološim znakom', *Godišnjak Filozofskog fakulteta u Novom Sadu*, vol. 4, pp. 151–63.

Ivić, M. (1961), 'On the structural characteristics of the Serbocroatian case system', *Internat. J. Slavic Linguistics and Poetics*, vol. 4, pp. 38–47.

Ivić, M. (1962), 'The grammatical category of non-omissible determiner', *Lingua* vol. 11, pp. 199–204.

Mirowicz, A. (1949), *O grupach syntaktycznych z przydowką*, Torun.

Potebnja, A. A. (1888), *Iz zapisok po russkoj grammatike*, vols. 1, 2, Xar'kov.

Trubetskoy, N. S. (1937), 'O pritjažatel'nyx prilagotel nyx (possessiva) starocerkovnoslavjanskogo jazyka', *Zbornik u Čast A Belića*, Beograd.

11 F. R. Palmer

'Sequence' and 'Order'

F. R. Palmer, '"Sequence" and "Order"', *Monograph Series on Languages and Linguistics*, vol. 17, 1964, pp. 123–30.

Both speech and the linguist's description of it are, or at least are often considered to be, unidimensional, the first in time, the second horizontally along the written page. The problem to be discussed here is, in its simplest terms, 'How far if at all, ought this shared feature of unidimensionality to determine linguistic analysis?' To distinguish between the single time dimension of speech and the single horizontal dimension of linguistic structure, J. R. Firth used the terms 'sequence' and 'order'. The terms are not clearly defined by Firth; the relevant quotations are

> In these structures one recognises the place and order of the categories. This, however, is very different from the successivity of bits and pieces in a undimensional time sequence (Firth, 1957, p. 5),

and

> Elements of structure, especially in grammatical relations, share a mutual expectancy in an order which is not a sequence (Firth, 1957, p. 17).

A more detailed statement, which I believe to be a fair interpretation of Firth is to be found in M. A. K. Halliday:

> Structure is made up of elements which are graphically represented as being in linear progression; but the theoretical relation among them is one of 'order'. Order may, but does not necessarily have as its realisation, sequence, the formal relation carried by linear progression; sequence is at a lower degree of abstraction than order, and is one possible formal exponent of it. A structure is thus an arrangement of elements ordered in 'places'. Places are distinguished by order alone; a structure XXX consists of three places (1961, pp. 254–5).

SEQUENCE, then, is of the observable speech events, ORDER of the linguist's constructs. The problem is to decide to what extent sequence is, or should be, an EXPONENT of order (I use, with Firth, the term 'exponent' for what Trubetskoy called 'realization' and Pike 'manifestation' – Halliday's use is a little different). I am not suggesting that all features of speech in their most directly observable form are in sequence. It is merely that we are here concerned with those features that ARE sequential. For there can be no denying that in the utterance 'cat' the velar plosive precedes alveolar plosive and in *John hit Bill*, *John* precedes *Bill*. Nor, of course, am I unaware of the fact that not all linguistic statements are made in terms of an 'ordered' (linear) structure.

Two extreme views are obviously possible on the question 'What is the relation between sequence and order?'. First, we may say that order is wholly determined by sequence, the order of the linguistic elements is identical with the sequence of those parts of the speech events they are intended to characterize. This is implicit in what C. F. Hockett has called the 'size-level' view of language, in which elements at each level occupy the same relative position as clusters of lower level elements of which they are composed (1955, p. 15). Secondly, we may say that order is wholly independent of sequence; if sequential features are relevant to the linguistic statement they are characterized by some means other than order. The point of view here is taken to its logical extreme by Hjelmslev and Uldall who say,

> in glossematic, as in logical algebra, ab is the same as ba . . . and the order in which functives are written should not be taken as having any significance (1957, p. 45).

Unfortunately this is never developed; they go on,

> since . . . order is utilized as a distinguishing feature in the linguistic expression . . . it is obviously necessary to devise some means of differentiating, e.g. *fist* and *fits*; but that is a problem which does not concern us here and which will be dealt with later.

But it is never dealt with.

The first view, though once firmly held, cannot be justified if we wish to make a statement in terms of ordered structure about any of the following characteristics of a language.

1. Features such as discontinuity and conflation. The first implies that what is a single item in the structure has as its exponents two or more elements that are separated in sequence. Conversely the second implies that two or more items in the structure have a joint set of exponents that cannot be sequentially divided.

2. Non-contrastive variations of sequence. If we do not wish to indicate these variations by a variety of structures we may set up one structure only, but clearly the order of the elements of this one structure cannot correspond to the variety of sequences. Halliday, for instance, considers *The man came(,) from the Gas Board* where *the man . . . from the Gas Board* is a nominal group, contrasting with (and differing intonationally from) *The man came from the Gas Board* where *from the Gas Board* is an adjunct (1961, p. 255, footnote 33). The structure of the first is SP, of the second SPA.

Unless I completely misunderstand him he is asserting that the first sentence is structurally the same as *The man from the Gas Board came.* It then follows that in no structural statement can the order of the elements correspond to both sequences.

3. Contrastive variations of sequence. It is possible to handle these together in a single ordered structure, but in that case it again follows that the order of the elements cannot correspond to the varying sequences. An example is to be found in Hockett when he suggests that *The body was found by a troop of boy scouts* may be regarded as *A troop of boy scouts found the body* and the morpheme {passive} (1961, pp. 51–2). Clearly the structure cannot correspond to both sequences (and the morpheme {passive} has no corresponding element in sequence).

We have good reasons then for not trying to handle everything in terms of linear structure. 2 and 3 are indeed a justification for transformational analysis, but I will return to this later.

The second view, which disassociates order from sequence has, as far as I know, never been put fully into practice. It is obviously uneconomical. If our linguistic symbols are written in order, it is convenient to make order significant, and the simplest use is to indicate sequence. A different use of order is found in the Arabic numeral system; the Arabs realized, though the Greeks did not, that order alone could distinguish units, tens and hundreds. If order is not made significant, moreover, some arbitrary rule will be needed, so that we shall not indiscriminately write the same symbols in various orders, for this would create confusion. An arbitrary rule of this kind is found in chemistry where H_2O and not O_2H is written though nothing different could be indicated by the latter. But there is no point in having an arbitrary rule of this kind, if a non-arbitrary (and useful) rule is available. Yet there may be parts of the linguistic description where the disassociation of sequence and order is forced upon us.

If more than one linguistic level is recognized, we may, of course, make different decisions concerning the relations of sequence and order at each level. In phonology, for instance, we may decide that order is wholly determined by sequence, while in morphology it may be wholly independent of it. In that case, it will follow that there is a difference between the order of the elements at each of the two levels. When considering grammar, then, we may handle the problem not in terms of relations between order and sequence, but of relations between the order of the grammatical elements. But it must not be assumed that there are no sequence/order problems in phonology. On the contrary there are problems of discontinuity, conflation and non-contrastive variation.[1] Contrastive variation is, as far as I know, not attested, but it could happen in theory; it might be convenient in a language to treat forms such as /pæt/ as an inverted form of /tæp/, /bæk/ of /kæb/ , and so on.

---

[1] For a contrary view see Martinet (1960, p.2).

Because of these difficulties, those linguistic theories that are in terms of linear structure compromise between the two extreme views. Order is TO SOME DEGREE determined by sequence. But there are two dangers here, one of saying too little about the relation of sequence and order, the other of saying too much.

First, there is the danger of not stating clearly when order is indicative of sequence and when it is not. Hockett has insisted that morphemes are not COMPOSED of phonemes but are REPRESENTED by them (1955, p. 15). It would have been consistent for him to continue with the statement that the order of the morphemes is not identical with the order of the phonemes (and therefore not with the sequence of the linguistic events), but that even between order at one level and order at another the relation is a representational one. In fact his notion of representation makes such a view inevitable. The morphemes of *took* and *liked* clearly cannot both be in the same order as the phonemes represented by them. It is worth contrasting here a consistent size-level interpretation such as that of B. Bloch (1947). Bloch interprets *took* as /tuk/+/Ø/. Such an interpretation allows for the morphemes of both *took* and *liked* to be in the same order as their constituent phonemes.

The second danger is a more serious one. It is, or should be, fairly obvious that if order is made indicative of sequence, we cannot, without confusion or redundancy, further establish a linguistic element whose exponent is sequence. Yet this is done. Z. S. Harris for instance, talks about morphemic order, even of 'morphemes of order' (1951, pp. 185-6; 1952, p. 18). Between the sentences *You saw Fred* and *Fred saw you* there is for Harris a difference consisting of a morphemic element. But these two sentences are not at all closely related linguistically. All that needs to be said to characterize them is that there is a structure in English which we may characterize NVN, and that *Fred* and *you* are both members of class N. The sequential features are taken care of by writing the two Ns in different places in the structure; in a similar way the difference between /kæt/ and /tæk/ is handled by writing CVC, and noting that both (t) and (k) are members of C. There is not something called 'morphemic order' or 'phonemic order' over and above the statement in terms of the elements N and C and their class membership. Similarly, of course, there is no important mathematical relation between 14 and 41, except that the same symbols are used in a different order. Moreover the fact that order is significant implies that the same symbols ARE being used in different places in the structure.

The confusion is perhaps seen more clearly still in Halliday, who has elements of structure S, P and O – Subject, Predicator and Object; one possible structure is SPO (1961, pp. 257–8). But a 'crucial criterion' of S (and presumably of O) is its place in sequence. The mere fact, therefore, of writing the symbols in an order is sufficient for their differentiation. Nothing is gained and clarity is lost by writing DIFFERENT symbols; instead of S and O we need one symbol placed in the first and the third position. Because he distinguishes between S and O, Halliday is compelled to add yet another kind of distinctive symbolization. He writes arrows over the symbols S and P ($\overrightarrow{SP}$) in order to indicate that sequence is an exponent of S, but it is only because he has distinguished between S and O that he now needs the arrows – to show that, except in terms of sequence, they are not different at all. If different symbols had not been used the arrows would not be necessary. We have then a double redundancy here, the second made necessary by the confusion created by the first. Of course, it may be argued that the justification for writing the separate symbols S and O is that it is not merely sequence that distinguishes them, but that there are the additional exponents of concord with the verb and of the form of the pronoun, and, indeed, that these are more important than sequence. For if we ignore intonation and juncture we can identify the subject and object of *Lots of money has she*. But if we treat concord, pronominal form and sequential position as exponents, we no longer require order. We may write SPO, SOP, PSO, etc., without distinction. Here, then, at this level, perhaps we should disassociate sequence and order. Why, then, does Halliday need two different symbols? There is a strong suspicion that S and O are not structural features of English, but some kind of universal (and notional?) feature of language. For he compares the English pattern in terms of S, P, and O with that of Latin; in Latin, of course, sequence is not an exponent of the distinction between the two.

Similar considerations hold for K. L. Pike's tagmemes, as long as a tagmeme is defined in terms of a particular place (slot) in a particular structure (1963). Yet some kind of tagmemic statement is essential if we wish to identify places in different structures, if for example (to use wholly traditional terminology) we wish to identify the object of an active verb with the subject of a passive one, or the subject of the first with the 'agent' of the second. Such identification is legitimate and indeed forced upon us if we undertake a transformational statement. We can use a set of symbols with subscripts, for instance, N_1 and N_2 such that *John* is N_1, and *Bill* N_2 in both *John saw Bill* and *Bill was seen by John*, but since it is customary in linguistics to use notionally informative labels for formally defined categories, terms such as 'actor' and 'goal' seem more appropriate.

Transformational analysis might seem to offer a solution (perhaps THE solution) to all our problems. Ought not all that part of linguistic description that can be dealt with in terms of an order dictated by sequence be handled in the constituent structure rules and the remainder in the transformation rules?

In fact, this simple solution is not adopted and perhaps it cannot be adopted. For categories that involve sequence are included in the constituent structure section and the sequential problems are handled in the morphophonemic section of the grammar. An obvious example in the English progressive will be the rule Prog. → BE+ING and a morphophonemic rule that transfers ING to the end of the verb.

The basic difficulty is that sequential and non-sequential features of language are often exponents of closely related formal categories (Bazell, 1953, p. 57). In English, for example, negation, if formally defined in terms of the /nt/ verbal formation, is closely connected with interrogation, if this is defined in terms of the change in the order of the subject and the verb. For both categories involve the auxiliary verbs. But while the former can be handled with no more difficulty (perhaps even less) in terms of linear structure than the past tense forms of the verb, the latter must be in terms of contrastive variations of sequence. (I do not think that the problem is seriously complicated by the introduction of DO since this can be treated morphophonemically as zero except where associated with /nt/.) Ought not, then, negation to be dealt with in the constituent structure rules and interrogation in the T-rules? N. Chomsky does not do this, but recognizes the structural relations between negation and interrogation and handles them together (1957, pp. 61–5). But this means that he is forced to treat them both in the T-rules in spite of the fact that negation would so much more naturally belong to the constituent structure section of the grammar. R. B. Lees on the other hand treats negation in constituent structure but interrogation in the T-rules and so fails to indicate the close relationship between the two categories (1960, pp. 19, 36).

Similar considerations hold for voice in English, for active and passive, *vis-à-vis* the other verbal categories of tense, progressive and perfect. All of these are formed in English by means of the auxiliaries BE and HAVE. A grammar that brings out the structural relationships between them ought to handle them all together, but voice involves sequential relations of a much more complicated kind than those associated with the other categories. All four categories could, of course, be handled together either in the constituent structure rules or in the T-rules, but the solution offered by transformational grammars simply fails to handle them together at all; voice goes into the T-rules and the other categories into the constituent structure rules (Lees, 1960; Fillmore, 1963, pp. 225, 228).

State Institute of Education West Bengal

One can therefore have a great deal of sympathy with Hockett when he suggests that passive could be treated as a morpheme in English. It is easy enough, of course, to show that this would greatly complicate the morphophonemic statement and this is obviously the reason why transformational grammars do not handle it in the constituent structure section. But if one talks, as Chomsky does, of 'deep underlying regularity' (1957 p. 68), or as P. M. Postal does of 'accounting for formal structures' (1964, p. 74), it is surely necessary that passive should be handled together with past, perfect and progressives. One can argue with some conviction that these categories differ only in the nature of their exponents.

Leaving aside the entirely different problem of embedding transformations, we may well ask what in principle distinguishes constituent structure rules and T-rules if the order/sequence relation does not. It seems that the distinction relies upon the WORD in that morphophonemic rules involving sequence are permitted ONLY where there is change of sequence within the WORD (as with negative, progressive and perfect), and not between words as with passive and interrogation. The distinction then rests upon a definition of the WORD or its acceptance as given in its traditional form. The usual explanation is that in constituent structure rules we are limited to the rewriting of one symbol at a time. But this does not help make the distinction. For we can easily handle passive interrogation, etc., in this way merely by leaving the more difficult order/sequence problem to the morphophonemics. A different point is made by Lees, that the constituent structure section allows us to formalize the notion of grammatical category (1960, p. 2); but this too is not a satisfactory explanation since passive and negative are as much grammatical categories as past, perfect and progressive.

Transformational grammar, then, disguises the problem of the relation between sequence and order. It does not answer it. The problem is still there though it may have been banished to the unexplored field of what C. E. Bazell has called 'pregrammatical statement', which is 'neutral with regard to different grammatical statements and indeed basic to . . . them' (1958, p. 19). Perhaps we need a pre-grammatical statement in which order is utterly divorced from sequence. Here and here alone, moreover, may we make 'God's truth' judgements about what are 'correct' and 'incorrect' interpretations of structure (in the widest sense), leaving us free for the sake of 'simplicity', 'economy', etc. to make 'hocus-pocus' statements based upon this interpretation in the grammar proper. For there is no reason to suppose that the criteria of simplicity and correctness will converge. Perhaps we should say, then, that it is only in the grammatical statement in this very restricted sense (where we are concerned with the purely practical task of stating formulaically our findings), that the problem of 'To what extent is order determined by sequence?' should arise.

References

Bazell, C. E. (1953), *Linguistic Form*, Istanbul.

Bazell, C. E. (1958), *Linguistic Typology*, School of Oriental and African Studies, University of London Press.

Bloch, B. (1947), 'English verb inflection', *Language*, vol. 23, pp. 399–418.

Chomsky, N. (1957), *Syntactic Structures*, The Hague, pp. 61–5.

Fillmore, C. J. (1963), 'The position of embedding transformations in a grammar', *Word*, vol. 19, no. 2, pp. 208–231.

Firth, J. R. (1957), 'A synopsis of linguistic theory', *Stud. Linguistic Anal.*, special volume of the Philological society, Oxford, pp. 1–32.

Halliday, M. A. K. (1961), 'Categories of the theory of grammar', *Word*, vol. 17, 3, pp. 241–92.

Harris, Z. S. (1951), *Methods in Structural Linguistics*, Chicago University Press, pp. 185–6.

Harris, Z. S. (1952), 'Discourse analysis', *Language*, vol. 28, 18–23.

Hjelmslev, L., and Uldall, H. J. (1957), *Outline of Glossematics*, Copenhagen.

Hockett, C. F. (1955), *A Manual of Phonology*, *Int. J. Amer. Linguistics* Memoir no. 11.

Hockett, C. F. (1961), 'Linguistic elements and their relations', *Language*, vol. 37, no. 1, pp. 29–53.

Lees, R. B. (1960), *The Grammar of English Nominalizations*, publication no. 12 of the Indiana University Research Center in Anthropology, Folklore and Linguistics.

Martinet, A. (1960), 'Elements of a functional syntax', *Word*, vol. 16, pp. 1–10. (See Reading 13.)

Pike, K. L. (1963), 'A syntactic paradigm', *Language*, vol. 39, no. 2, pp. 216–30. (See Reading 14.)

Postal, P. (1964), *Constituent Structure: A Study of Contemporary Models of Syntactic Description*, publication no. 30 of the Indiana University Research Center in Anthropology, Folklore and Linguistics, p. 74.

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road

Part Three
Relations, Functions and Structures

Here we find representatives of other schools or viewpoints: Harris (Reading 12), the leader of his own school and the preceptor of Chomsky; Martinet (Reading 13), another leader; Hays (Reading 16) presenting Tesnière's theory; Halliday (Reading 18) his own; and Reich (Reading 19) presenting Sydney Lamb's Stratificational Grammar. Pike here speaks for himself (Reading 14), but is also supported by Healey, who offers some fascinating facts concerning an illocutionary marker in a New Guinea language (Reading 15). New Guinea seems to be the last frontier for linguistic discovery, a role that in previous generations has been played by Far Eastern languages, by Oceanic and Uralic-Altaic, by African, by Caucasian and American Indian (North, Central and finally South), and striking novelties in phonology as well as syntax are appearing.

12 Z. S. Harris

Co-occurrence and Transformation in Linguistic Structure[1]

Excerpts from Z. S. Harris, 'Co-occurrence and transformation in linguistic structure', *Language*, vol. 33, 1957, pp. 283–340.

1 The method of analysis

1.1 Individual co-occurrence

The range of individual co-occurrence of a morpheme (or word) i is defined first of all as the environment of morphemes (or words) which occur in the same sentences with i (in some body of linguistic material). This is indeed the initial information available for morphological structure. Given utterances or sentences and given the morphemes, we can say that each morpheme i has the particular morphemes j, k, . . . as co-occurrents in the sentences in which i appears. Each morpheme has a unique set of co-occurrents (except for special morphemes such as some paradigmatic affixes which all occur with the same set of words and in the same sentences).

The individuality of morphemes in this respect makes it difficult to set up any compact description of a language. However, some morphemes have very similar (though not identical) sets of co-occurrents: thus, the set of co-occurrents for *cloth* – e.g. *The* ( ) *tore*, *The* ( ) *was torn*, *Get me a* ( ) *quick* – may have many morphemes in common with the set for *paper*, certainly many more than with the set for *diminish*. This suggests that morphemes can be grouped into classes in such a way that members of a class have rather similar sets of co-occurrents, and each class in turn occurs with specific other classes to make a sentence structure. In structural linguistics this classification is not set up on the basis of relative similarity of co-occurrents, but rather on the basis of a particular choice of diagnostic co-occurrents: *cloth* and *paper* both occur, say, in the environment *the*

---

[1] The study of transformations arose out of the attempt to construct a method for analysing language samples longer than a sentence; a preliminary list of English transformations is given in Harris (1952). Although the present paper was presented at the 1955 meeting of the Linguistic Society in Chicago, some of the details on English transformations have been added since that time, in connection with a project of the National Science Foundation.

From a time when this work was still at an early stage, Noam Chomsky has been carrying out partly related studies of transformations and their position in linguistic analysis (1955; 1956; 1957). My many conversations with Chomsky have sharpened the work presented here, in addition to being a great pleasure in themselves.

( ) *is* (i.e. after *the* and before *is*), where *diminish* does not appear; we call this class N. And *diminish* and *grow* both occur, say, in *It will* ( ), where *paper* and *cloth* do not; we call this class V. The diagnostic environments (stretches of co-occurrents) are chosen in such a way that the resulting classes permit compact statements about co-occurrence. For example, *cloth*, *paper*, *diminish*, *grow* all show some differences in their environments, so that no simple summary can be made. But in terms of the classes N and V we can say that every N occurs before some V in the environment *the* ( ) V, and every V occurs in the environment *the* N ( ) for some N.

When we proceed to describe the structure of sentences (i.e. the choices of morphemes that occur in a sentence) in terms of these classes, we find that the work is of manageable proportions. In the sentences of any given language, only certain sequences of classes (set up as above for that language) will be found; and these sequences which constitute the sentences can be described as the products of a small number of elementary class sequences (constructions) which are combined in certain stateable ways; for instance, the class sequence TNPNV (*The fear of war grew*) results from the elementary class sequences TNV and NPN by the substitution operation NPN = N.[2] This compact description of sentence structure in terms of sequences of classes is obtained, however, at a cost: statements such as NPN = N or 'TNV is a sentence structure' do not mean that all members of NPN have been found in the same environments as all members of N, or that all sequences of members of T, N and V in that order have been found as sentences. [. . .]

1.2 Constructional status

Although, as we have seen, individual co-occurrence cannot be used directly in discovering morphological relations, we do not have to disregard it (as we do in structural linguistics). In spite of the impossibility of obtaining complete or definite co-occurrence data, we can still consider various indirect questions (absolute or relative), such as whether a particular morpheme or class has some special property in its co-occurrence set, or how the co-occurrence sets of two classes or class sequences compare. Such questions can fit into the class structure because each co-occurrence check would be within stated classes and constructions (by the adjusted definition); and the results can be then phrased as new statements, again about classes and constructions, as will be seen below. From the point of view of structural linguistics, this amounts to asking one of the few types of outside questions that are still relevant to it, for these are questions which are couched in terms of the raw data of structural morphology (the occurrence of morphemes in sentences), and which lead to additional information about inter-class relations, yet which had not been asked in the original study of class environments.

---

[2] Notation: Morpheme and word classes: N (noun) and V (verb) as above; A (adjective) includes *large*, *old*, *extreme*, etc.; T (article) includes *the*, *a*; P (preposition) includes *of*, *from*, *to*, etc.,; C (conjunction) includes *and*, *or*, *but*, etc.; D (adverb) includes *very*, *well*, *quickly*, etc. Classes of affixes (mostly suffixes, some prefixes): na indicates an affix such that N+na (i.e. the sequence N na) yields an A word (a word substitutable for A morphemes): *papery*, *cloth-like*. Similarly vn after V yields an N word: *growth*; nn after N yields an N word: *growths*, *childhood*; and so on. And v (tense and verb 'auxiliary' class, para. 2.5, p. 158) includes *-ed will*, *can*, etc. S stands for sentence. The equivalence A = B indicates that if sequence A occurs in a sentence, the substitution of B for A also yields a sentence. All these classes apply specifically to English; the analyses and transformations below have been set up for English data, although the principles are general and not limited to any particular language.

Individual co-occurrences can be expressed in structural terms by saying that they are the values for which a structural formula is satisfied. That is, if we regard each statement about constructions (NPN = N, or TNV = S) as a formula which holds when particular morphemes (or sequences of them) occur in the positions indicated by their class marks, then we can say that the particular combinations of morphemes which co-occur in various instances of a construction are those which satisfy it (i.e. those for which the formula holds).

We consider now the major applications of individual co-occurrence data within constructional morphology.

Given a construction (which is recognized by its place within larger constructions, up to a sentence), we can see some relations or different statuses among the participating classes by noting details of class occurrence for each class. For example, given the AN construction (*slight hopes*), we note that it is substitutable for N alone but not for A alone: both *slight hopes* and *hopes* occur in *Their* ( ) *faded.* We can express the constructional equivalence AN = N by defining in the place of these two a composite construction (A)N, with the parentheses indicating occasional omission; and in this new construction we can say that N is the 'head', meaning that it always is present when the construction occurs. Another kind of differential status for the classes in a construction is seen when one class has a correlation between its subclasses and some other element in or out of the construction (grammatical agreement); e.g. when nouns have each a particular gender suffix while the neighboring adjective takes whichever gender suffix the noun has. (Again we might say that the noun is the 'head'.) [. . .]

All these considerations together (the position, co-occurrence ranges, and other properties of classes or subconstructions within a construction) determine what we may call the constructional status or relation of each participating class or construction in a (larger) construction. We find various types of status in respect to individual co-occurrence: in Vv neither class has restricted co-occurrences (each v_i occurs with all V, and each V_i with all v), but in VP (*chalk up*, *tie up*, *tide over*) both have. In AN each class has fixed co-occurrences in the other, but only N has fixed co-occurrences in the neighboring V. The types of constructional status are useful in characterizing each construction, and in analysing various less obvious construction (para. 2, p. 157). Many of them serve to explicate intuitively known grammatical relations: in the AN type one says that A modifies N. Finally, the types of constructional status themselves are partly determined by investigable conditions, such as the size of the participating classes.

Aside from the matter of constructional status there are certain subclasses of the major classes which are characterized only by peculiarities of their individual co-occurrence (para. 2.6, p. 160).

1.3 Transformations defined

In addition to investigating the types of co-occurrence for various classes in a construction, as in para. 1.2, we can compare the co-occurrences in two different constructions which contain the same classes. In many such constructions, as in N'*s* N and N *is a* N above, the co-occurrences are different. In some constructions the co-occurrences are about the same, and it is for these that transformations will be defined.

If two or more constructions (or sequences of constructions) which contain the same n classes (whatever else they may contain) occur with the same n-tuples of members of these classes in the same sentence environment (see p. 152), we say that the constructions are transforms of each other, and that each may be derived from any other of them by a particular transformation. For example, the constructions N V v N (a sentence) and N'*s* V*ing* N (a noun phrase) are satisfied by the same triples of N, V and N (*he*, *meet*, *we*; *foreman*, *put up*, *list*; etc.); so that any choice of members which we find in the sentence we also find in the noun phrase and *vice versa*: *He met us, his meeting us* . . .; *The foreman put the list up*, *the foreman's putting the list up* . . . Where the class members are identical in the two or more constructions we have a reversible transformation, and may write e.g. N_1 v V $N_2 \leftrightarrow N_1$'*s* V*ing* N_2 (and the set of triples for the first = the set for the second).[3] [. . .]

Since transformations are based not on the absolute contents of a set of co-occurring n-tuples, but on the similarity between the sets for different constructions, we do not even require a yes-or-no decision on whether any given n-tuple is in the set. We can accept graded decisions, and we do not even have to assert that any particular n-tuple never occurs. For example, the formulation might be that to the extent that any n-tuple satisfies N_1 v V N_2 it also satisfies N_1*'s* V*ing* N_2: *he*, *meet*, *we* certainly in both forms; *moon*, *eat*, *cheese* doubtfully or for particular environments in both forms; *soup*, *drink*, *abstraction* hardly at all in either form.

---

[3] Same subscript means same member of the class: the second appearance of N_1 indicates the same morpheme as the first N_1.

There may be differences in some respects between the sets satisfying two constructions, without precluding a transformational relation. For example, all or certain ones of the n-tuples may be less frequent or natural in one construction than in another. Differences of this kind need not restrict the transformation, if they can be indicated in an adequate manner.

As to the differences between the same-class constructions – for there must of course be differences between them – these may consist of difference in order, or in individual morphemes (affixes or words), or in added small or large classes. If it is impossible to specify which members of the added classes occur in each instance of the construction, the statement of the transformation remains incomplete [. . .]

The general requirement can be stated as follows: no domain (section of the sentence or the like) to which the two constructions are grammatically connected may admit co-occurring morphemes for one of the constructions which are not admitted for the other. Thus the adverbial position at the end of the examples above is connectable both to N V*ed* and to N *will* V (both N V*ed* D and N *will* V D are grammatical sentence formulas); therefore it would have to admit the same members in both cases, if these two are to be transforms. The effect of this requirement is to make the identity of co-occurrence stretch over the maximal grammatical domain, i.e. to preclude any inequality (restrictions) of co-occurrence that would narrow down what the grammar permits. In this sense the requirement safeguards the characteristic of transformations, which is to have no co-occurrence restrictions that are correlated with the difference between the two constructions (i.e. with zero as against *be . . . -en . . . by* plus reverse order in the active/passive, or with *-ed* as against *will* in the present rejected case). Note that this formulation does not contravene the transformational standing of N v V N and N'*s* V*ing* N, even though the sentence environment of the latter is lacking in the former. The latter, being a noun phrase, is always connected to the remaining constructions of its sentence (*We're uncertain about his meeting us*; *The foreman's putting the list up caused the wildcat*). But these additional sentence-parts cannot be connected to N v V N since that is a sentence in itself, hence their absence here is due to grammatical conditions and not to any inequality of co-occurrence.

The consideration of meaning mentioned above is relevant because some major element of meaning seems to be held constant under transformation.

There may be differences in meaning between two constructions that are transforms of each other, due to their external grammatical status: e.g. the fact that N v V N is a sentence while N'*s* V*ing* N is a noun phrase. There may be differences of emphasis or style, as between the active and the passive. And certain transforms have specific meaning differences associated with the specific morphemes they add, as between assertion and question. But aside from such differences, transforms seem to hold invariant what might be interpreted as the information content. [. . .]

1.4 Determining the evidence

To establish the transformations in any given language we need methods, and if possible an organized procedure, for seeking (para. 1.41) what constructions may contain identical co-occurrences; these methods should if possible be general, but additional ones may be based on special features of the language. And we need methods of checking (para. 1.42.) the co-occurrences in each construction, so as to see if they are indeed identical.

1.41. *Domains of transformation.* To find domains that may be transforms, we need to consider only constructions, since it appears that a given class has specifiable relations (restrictions) of individual co-occurrence only in respect to constructions, e.g. the relations appear between the head A class and non-head class of a construction (N and A respectively in AN), or between the head of construction c and the head of construction k when c+k in turn constitute a larger construction (N head of the noun phrase and V head of the verb phrase in the sentence T A N+V v D, *The new girl laughed loudly*). There are no co-occurrence relations between the A and the D here, these being two non-heads in the two constructions; but there is such a relation within the adjective phrase DA, where the A is the head (*partly new*). These domains may, however, be sequences of constructions, as when we wish to compare the sentence sequence $T\ N_1\ V_2.\ T\ N_1\ D\ V_3$ T N with $T\ N_1$ *who* $V_2\ D\ V_3\ T\ N_4$ (*The king abdicated, # The king soon resumed the throne* vs *The king who abdicated soon resumed the throne*). [. . .]

1.42. *Obtaining co-occurrence ranges.* To check whether the co-occurrences are identical in two n-class constructions (suspected of being transforms), we obtain a list of n-tuples which satisfy the first construction and ask about each one whether it also satisfies the second; and vice versa. Or we hold constant n — 1 members of an n-tuple, and ask for a list of members of the *n*th class satisfying the first construction and a list of members of that class satisfying the second. For example, to test N_1 v V N_2 and N_2 v *be* V*en by* N_1, we begin, say, with *The cat drank* ( ) and ask for a list of words that would fill out the sentence, and then we offer ( ) *was drunk by the cat* and ask for a list of words that would fill out that sentence.

2 Co-occurrence as a structural property

2.0

Just how much grammatical knowledge is needed before transformations can be investigated depends in part on how much work is put into discovering the transformations. It may be enough merely to identify the morphemes and to have a constructional analysis of the simpler sentence types (as in para. 5.4, p. 179). For the present formulation, however, we will assume the whole of the usual structural grammar of the language, in order not to have to distinguish parts which are not needed.

2.1 Dependent elements in constructions

While some parts of grammar may not be needed, it is helpful if the parts which are used (e.g. the morpheme classes and the major constructions) have a very detailed analysis; for then the transformations are more apparent. This means a detailed investigation of how the occurrence of one morpheme or sequence (or stress or tone) in a sentence depends on the occurrence of some other in the same sentence or nearby, so that a change in one is accompanied by a change in the other. For example, if we offer an informant the utterance *The letter was returned # It must have been misaddressed* and ask him to substitute *The letters* for *The letter*, it is almost certain that he will also substitute *They* for *It*. If we get as many tense substitutions as we can in the sentence *The dog bit me as if he had been a man* we will probably find some tense combinations lacking, indicating some dependence between the two tense morphemes in this position. Some elements are completely dependent, as the *my* (or the *I*) in *I saw myself*. [. . .]

2.4.C. Members of the class C (*and*, *but*, *more than*, etc.) have in general the property that they occur between two instances of the same construction. Given an utterance containing C, it is in general possible to find a construction X immediately before C and a construction Y immediately after C, such that X and Y have the same status within the next larger construction; examples are *a pale but cheerful face*: two A within an N phrase; *the low wages and long hours in Detroit*: two A N sequences within an N phrase; *The low wages and the long hours were the chief causes*: two N phrases within a sentence; *Either I go or you go*: two sentence structures within a sentence. This means that both X alone and Y alone can occur in the same structural environment as X C Y: *a pale face*, *a cheerful face*.[4]

[4] Where X (or Y) in the given environment participates in a discontinuous morpheme extending to another construction, we find this participation in both X and Y of X C Y: *He reads*, *He reads and writes*. Certain C may be defined as participating in such an extending morpheme even where their X or Y alone would not: in N_i *and* N_j V, the V generally shows the plural morpheme even when N_i V or N_j V alone does not (*A man and woman are here* but *A man is here*).

In some cases it seems as though we do not have identical constructions on both sides of C: *I'll take the first and you the second*; *We got there on time but not he*; *He came, and fast*; *He's bigger than you.* In para. 3.1 [not included here], however, it will be seen that the shorter construction in each case contains zero forms of morphemes which are visible only in the longer construction, and precisely enough zeros to make it the same construction.

Not only the class environments, but also the individual co-occurrences of X and of Y alone are often the same as for X C Y. This holds for certain members of C: *and*, many occurrences of *or*, etc. Even for *and* there are certain pairs of X Y which do not occur singly in the same individual environments in which they occur together: *She and I don't see eye to eye*; *Sugar and water make syrup*; *It rocked to and fro.* These pairs, or certain properties of them as a subclass, can be investigated, in order to distinguish them from all the others. There are great differences, both in this and other respects, among the various C.

2.5. v (*auxiliaries*). If we compare *He paints*, *He painted* with *He doesn't paint*, *He dóes paint*, *Did he paint?*, *Only then did he paint*, *I painted and so did he*, we see that in the presence of certain conditions (*not*, emphatic stress, question intonation, etc.) V+its suffix is replaced by *do*+that suffix+V; and that in a subset of these conditions (question, or after certain words) *do*+the suffix further changes places with the preceding N. If we consider that the main V here remains *paint*, it will not be easy to explain the occurrence of *do* here. If we consider that the main V in the altered sentences is *do*, with *paint* as some secondary element (as it is e.g. in *He learned to paint*) the grammar will be simpler. But it will turn out that the individual sentence environments in these altered sentences will be identical with those where *paint* is the V, even though in the altered sentences the V is now *do*. This is not the case in other sentences where *paint* is a secondary element: the individual sentence environments around *doesn't paint* are closely related to those around *paint*; but for, say, *learned to paint* we can find environments where *paint* might not occur, such as *He learned to paint from the impressionists.*

If now we compare *He will paint* with *He will not paint*, *He wíll paint*, *Will he paint?*, *Only then will he paint*, *I'll paint and so will he*, we see that under the same conditions the V+auxiliaries (*will*, *can*, etc.) do not change, except that in the subset of conditions mentioned above the auxiliary changes place with the preceding N. This and other considerations which will appear in para. 2.6, p. 160 suggest that the *-s*, *-ed* be considered affixes of *paint* even after they move in front of it, and that the *do* which precedes them be considered not a morpheme at all but only a phonemic carrier for the suffixes when they do not have their V before them. (The suffixes occur only after a phonemic word, and interchange in position with V leaves them without a phonemic word.)

On this basis we can now analyse the forms which have neither suffix nor auxiliary: *They paint*, *They do not paint*, *They dó paint*, *Do they paint? Only then do they paint*, *We paint and so do they*. We say that there is a zero suffix (variant of the *-s*) after the V *paint*, which moves in front of the V under the conditions indicated. Like the other suffixes, this zero is always attached to a preceding phonemic word, and here receives the carrier *do*.[5]

There are now two different interchanges of position ('inversions'). One is that the suffixes *-ed* and zero (with variant *-s*) of V appear in front of V under all the conditions of the type illustrated above; they then have the carrier *do*. The other is that both of these suffixes and also the auxiliaries *will*, *can*, etc. appear in front of the preceding (subject) N in the subset of conditions exemplified above. The first inversion simply brings the suffixes over to the auxiliary position, when certain elements (including the emphatic) come before the V.[6]

The class v has the following properties: (a) each occurs with each V; (b) V never occurs without some v or other, nor v without some V (apparent exceptions will be discussed in para. 2.63, p. 162); (c) each v may have some restrictions as to the separable D, PN in its sentence (e.g. *-ed* will hardly occur directly with *tomorrow*); (4) but aside from such restrictions, which apply to a given v equally for all V, the individual sentence environments of each V remain the same no matter what the v. Note that this is not the case for, say, *learned to paint*, *learned to talk*, where the sentence environments differ in various ways from those of *painted*, *talked*.

The first and third of these three properties also apply to the constructions in *He is painting*, *He has painted*, *He has been painting*. All constructions of this type can be described by saying that between the v and the V there occur, sometimes but not always, *have ( )en*, *be ( )ing*, or the first of these followed by the second; the empty parentheses mark the location of the following V word (including *have* and *be*). Then *He paints*, *He can paint* have v V; *is painting*, *may be painting* are v+*be* ( )*ing*+V; *has painted*, *had painted*, *will have painted* are v+*have* ( )*en*+V; *has been painting*, *could have been painting* are v+*have -en*+*be -ing*+V. These constructions may be viewed as expansions of v, constituting v-phrases of which the v proper is the head.[7] [. . .]

---

[5] This zero morpheme meets the requirements for a zero – of being determined by observable conditions (namely *I*, *you* and plurals as subject); and it satisfies the meaning of V+zero, which is not just V in general but V in the present tense. The *-s* 'third person singular' can now be considered an alternant of this zero, occurring with the remaining subjects.

[6] For convenience, we will henceforth consider the *-ed* and zero (and *-s*) suffixes to be members of the auxiliary class v; this requires us to view *could*, *should*, etc. as independent members of v rather than as *can*, *shall*, etc. + *-ed*. Alternatively we can regard the suffixes as a separate class, and *could* as *can* + *-ed*, etc. But *should*, *would* often occur in environments where *-ed* does not otherwise occur (e.g. *I would, if he will too*). Furthermore, we would have to note that *can* etc. occur with *-ed* (to make *could*) but not with *-s (He can*, but *He sees)*. Note in the whole analysis the departure of structural considerations from historical antecedents.

2.6 Pro-morphemes

There exist morphemes whose X-co-occurrents (for each class X in constructional relation to them), in each sentence, equal the X-co-occurrents of a morpheme (of class Y) occupying a stated position (or one of several stated positions), relative to them, in the same sentence (or sequence of sentences), and whose total X-co-occurrents in all the appearances of these morphemes equal the sum of the X-co-occurrents of all the members of the class Y (which occupies the stated position relative to them). Such morphemes will be called pro-morphemes of the class Y, or pro-Y. If the Y position with the same X-co-occurrents (roughly what is called the antecedent position) is uniquely determined, the pro-Y will be called bound. If it is not completely determined, the pro-Y may be called indeterminately bound; in this case we cannot tell definitely to which antecedent the pro-Y 'refers'. The pro-Y can be viewed in each sentence as a positional 'variant' form of its antecedent in that sentence.[8]

2.61. *Pronouns.* The morphemes *he*, *I*, etc. which appear as pro-morphemes of the class N usually occur indeterminately bound, so that one cannot specify without investigation which preceding or following N has the same range of co-occurrents: *Mark came too late*; *Carl had taken his*

---

[7] These are only combinations that constitute verb structures. To each of these can be added *be -en* just before the V, forming the passive of that construction; but this involves also a certain rearrangement of the sentence environment (para. 4.11, p. 170). A few other combinations occur, as in *He has had stolen from him two invaluable manuscripts*, which seems to contain v+*have -en*+*have -en*+V. But the constituent analysis here is not *he* (subject N)+*has had stolen* (V)+P N+object N-phrase, for these cannot be rearranged like other N, V, P N, or replaced here by other N, V, P N. This sentence has to be analysed as a transformation; even without that the *stolen* would be taken with *from him*, not with *has had*. Note the alternative form *He has had two invaluable manuscripts stolen from him.*

[8] In English the antecedent is almost always in a different sentence structure from the pro-morpheme, though both sentence structures may be within the same sentence. (Sentence structure is used here to mean any sequence which is a transform of a sentence, such as *which had cracked* in *The wall which had cracked finally collapsed.*) An exception is pronoun+*self*, whose antecedent is almost always the subject: *Carl saw himself*, *Carl went by himself*, *Carl himself saw it*. This can perhaps be analysed as not a pronoun but just a morpheme *self* (different from *the self*), with the subject having a discontinuous repetition (in pronoun form) before this *self*. However, pronoun+*'s* (which is a member of A) occurs freely in the same clause as its antecedent N: *He took his books.*

painting and gone off. In many cases this pronoun is not bound at all, that is, it does not have the co-occurrence range of (and does not refer to) any N in its neighborhood, as in a story in which a character is never referred to except as *he*. Such pro-morphemes may be called free; they lack the pro-morpheme property of having antecedents, but possess the other property of having a co-occurrence range equal to the sum of co-occurrence ranges of a stateable class (N). This is equivalent to saying that in each appearance of a pronoun, including a free pronoun, one can always find some noun of its subclass that could substitute for it, i.e. could be found in its position in that sentence.[9] [. . .]

All these are (with some individual differences) pro-morphemes of the full N-phrase (including N-phrase P N-phrase = N-phrase: *The book of old songs . . . It . . .*). This is a way of stating the fact that we do not find, for example, *the* or *old* before *it*, or *of songs* after *it*. These are therefore pro-N-phrases, and are members of the class N-phrase rather than of the class N. Hence *he*, *this*, etc. include T within them (i.e. they constitute a recurrence, or a variant form, of some neighboring N-phrase with its *the*, *a*, *some*, etc.), and this gives them an individuating semantic effect. That is, in *The man . . . he . . .*, the pro-morpheme *he* indicates a recurrence not merely of the morpheme *man* but of the particular *man* individuated by its article *the*. [. . .]

2.62. *The wh- pro-morphemes*. If we consider *who*,[10] *which*, *what*, *where*, *when*, *why*, etc., we find that they can all be described as occurring in a specific position, namely before a sentence structure S_2 which lacks some particular constituent. The methods of para. 2.1 (p. 157) show that the *wh*-word has precisely the status of this constituent.[11] Furthermore, since *wh*- is common to all these we extract it as a morpheme which indicates this construction as a whole, while the residues of each *wh*- word are morphemes each with the specific constructional status of the section that is lacking in S_2. The *wh-* itself then precedes a complete S_2 (made complete by inclusion of the post *-wh-* element). In the thus completed S_2, *-o* is subject N-phrase, *-om* object N, *-ose* is N'*s* (which is a member of A), *-ich* and *-at* are N or A (A if an N follows, as in *which books*, *what books*), *-ere*, *-en*, *-ither*, etc. are P N, and *-y* (of *why*) can be taken as P N or C S (subordinating conjunction plus sentence).

---

[9] The total co-occurrence range of a free *he* equals the total co-occurrence ranges of all bound *he*. A marginal case is such pronoun occurrences as unstressed *you're*, *we're*, *one's* in ( ) *only young once*, which can perhaps be replaced by all members of N, or perhaps only by a new subclass of N (*people*, *Jim*, etc.; whereas *one* can otherwise be substituted by any N). In contrast, the *it* of *It's raining* is not a pro-morpheme at all: almost no other morpheme can substitute for it here, so that this *it* is just a member of N.

[10] The following analysis will show *who* to contain the morpheme *wh*- even though it lacks the phonemic /hw/.

[11] Various additional considerations show that these morphemes belong to the classes missing in the sentence structure that follows them. For example, P occurs in general only before N, but it also occurs before *whom*, *which*, *what* (*to whom*, *from which*, etc.). Throughout para. 2.62 N indicates N-phrase: the *wh*- words replace a full N-phrase and do not have T or A before them; and V will be used for V-phrase. Whenever S is written here directly after *wh*- it is understood that some part of this S is occupied by the post-*wh*- morpheme, e.g. . . . *who went* is *wh*- S, the *-o* being the subject N; *whom I saw* is *wh*- S, the *-om* being the object N.

There are three main positions in which the whole *wh-*$+S_2$ occurs: with question intonation, as adjective-phrase after nouns, and as object or subject of another sentence. The list of post-*wh-* morphemes is slightly different after each of these. [. . .]

2.63. *Pro-V*. If we compare *I'll go if you will* with *I'll go if you will go*, we see that *go* can have a zero variant in its second occurrence; this zero is analogous to, say, *-o* as second-occurrence variant of *the man* in *the man who*, and satisfies the conditions for being called a pro-verb. In all cases like the short form above, the one V (*go*) is in the V-co-occurrence for both N (*I*, *you*), and is the antecedent of the bound zero pro-morpheme at the end. The auxiliary v may be different (*I'll go if you can't*) or the same, but in any case it occurs in full in the second part, so that the pro-morpheme covers only the V and not v+V (there is no *I'll go if you*). When we consider *I left when he did*, we understand (on the basis of para. 2.5, p. 158) that this is the same construction: after the second N there is no V (*did*) but only the *-ed* tense (member of v) with a carrier *do*. The pro-V here is not *do* but zero, as in the other members of v. We find the tenses and other auxiliaries intermingled in *I left so that he would*, *I can go if they did*. We therefore analyse *I can* not as N+V but as N+v+zero pro-V, the antecedent V being usually present in the neighborhood. This accords with the absence after *can* and the like of the third-person *-s* which is found after all V.

The zero is actually a pro-morpheme of the V-phrase, including V P constructions, the object N and V+P N ('indirect object'): *I'll go up if you will*, *He likes the nineteenth-century writers but I don't*. It is possible to have a pro-V with an object different from the antecedent; but in that case the auxiliary is included and we have a pro-vV: *I got the first copy and he the second.*[12] [. . .]

---

[12] After constructions like *I am going*, *I have gone*, the second part is *but he isn't*, *but he hasn't*, not *but he doesn't*. The *is* and *has* operate like members of v. After *I will be going* the second part is *and so will you be* or *and so will you*, with the *be* repeated like a member of v or zeroed like a V. A similar situation occurs in N *is* N sentences: *He's a fool but she isn't*, *He'll be late but she won't be* or *but she won't*. And the case is similar for V *to* V and V V*ing* sequences: *I'll try to catch it if you will* or *if you try*.

2.8 V N N constructions

Two chief types of constructions seem to have two object N after the V. In one of these, e.g. *give the fellow a book*, *ask him a question*, the first N is always one which also appears in the P N range of that V (*give to the fellow*, *ask of him*); the second N may be one which is not in that range (*give tone to a book*, but hardly *ask P a question*).[13] Only a small number of V, the subclass V^a, occur with such two object N.

A very different group is seen in *The oil magnates made Harding president*, *The committee named him an honorary member*, *We found the barn a shambles*, etc. The V here, in general different from V^a, comprise a subclass V^b. Neither N is necessarily in the P N range of the V. The distinguishing property is that the two N are always in each other's co-occurrence range (in the same order) for the construction N_1 V^n N_2 (where V^n is a V subclass including *be*, *become*, *seem*, *remain*, etc.). That is, we can find sentences *Harding is president*, *He became an honorary member*, *The barn remained a shambles*. Furthermore the two N participate jointly in discontinuous morphemes which extend beyond a structure, such as *-s* 'plural': *He named them honorary members*.[14] [. . .]

2.9 Summary of constructions

A detailed analysis of the class constructions within each sentence type may make it possible to say that, of two sentence types A and B, A has the same construction as B except for the addition of some x: $A = B+x$. For example, *He didn't go* = *He went* + *not* (para. 2.5, p. 158). When we add to this a consideration of the actual word choices (individual co-occurrences) we may be able to show that a single n-class construction is satisfied by several different sets of n-tuples, sets which occur elsewhere in different constructions. For example, N P N is satisfied by one set of triples which also occur in reverse order (N_2 P N_1), and by another set in which the two N may each occur alone in the position of the N P N, and so on (para. 2.3 [not included here]). When we compare various partially-similar constructions we may find that one of them can be considered a special case of another. For example, *who went* can be considered a special case of an N V sentence (like *Carl went*) on the basis of considerations which make *-o* a special member of N (para. 2.1, p. 157; para. 2.6, p. 160). And when we compare the word choices in partially similar sentences or in neighboring sections of a sentence (para. 2.1, p. 157), we may find that undecidable or unique constructions can be reinterpreted (by environmental considerations) as special cases of known constructions, as when *He knows more than I* is shown to be *He knows + more than + I know*. By so comparing sentences with similar or neighboring sentences, we prepare the ground for showing that all sentence structures are combinations or transformations of just a few simple sentence structures. [. . .]

---

[13] It will be noted later that for each such construction V N_1 N_2 the same morphemes also occur in the construction V N_2 P N_1: *give a book to the fellow*, *ask a question of him*. But this is a transformational criterion, such as we are skirting here: the constructional status here is to be determined only by grammar and co-occurrence ranges. Some additional transformational properties characterize only this structure: the transformation of replacing the second N by its pro-N generally involves interchange of the two N (*give it to the fellow*) but not so for the first N (*give him a book*); there are two passive transformations, one of which shows the P N character of the first N (*A book was given to the fellow*, *The fellow was given a book*).

[14] Similarly in N V^n N: *They are honorary members*. The *shambles* above does not conflict since it is preceded by the singular *a* and is hence singular like its first N *barn*.

3 Transformations in sentence sequences

3.0

Certain sequences of two or more sentences have a special form for one of the sentences (usually the second): e.g. *Some groups have rebelled frequently # Some only rarely*, or *I've just been over there # Oh, you were there?* In addition, many sentences which have what might be called complex structures can be analysed as containing a sequence of two or more sentences or sentence structures, some or all of which have special forms: *I met him coming back*. In all these cases the sections with special forms (*Oh, you were there?* or *him coming back*) can be shown to be transforms of ordinary independent sentences, in the sense of the definition of para. 1.3 (p. 154). These transforms can thus be viewed as variant forms of sentences. Some of these variant forms are positionally bound, occurring only in particular sentence sequences, e.g. *Some only rarely* or *him coming back*. Others also occur outside of the sentence sequences, e.g. *Oh, you were there?* [. . .]

Finally, the fact that these are indeed transforms will be discussed in 4, where both these and the other transformations of English will be listed.

The result of 3, then, will be to show that various types of sentence sequence and of complicated sentences are the product of one sentence with the transform of another sentence. There are certain transformations which so to speak change a sentence into a noun-phrase or an adjectival phrase or a subordinate or coordinate clause. That is, given sentence S_2, such a transform of it $T\,S_2$ occurs next to another sentence S_1, with $T\,S_2$ filling one of these positions (N-phrase, etc.) within S_1. There is no correlation in English between the transformations and the positions they fill: several different transformations – $T_1\,S_2$, $T_2\,S_2$, etc. – may fill the same position within S_1 (e.g. there are several ways in which S_2 may be nominalized and appear as object of S_1); and the same transformation $T_1\,S_2$ may fill different positions in S_1. Quite apart from its transformational form, there are restrictions of individual co-occurrence dictating which S_2 will fill which positions in S_1: not all S_1 S_2 combinations occur. [. . .]

3.2 *Types of sentence combination*

Establishing the antecedents of pro-morphemes, and placing zero morphemes within sentence combinations, can show that these sequences contain two or more sections, each of which has a sentence-like internal structure. It remains to investigate the different types of combination.

In some of these combinations what we may call the secondary section becomes a full sentence as soon as the pro-morphemes and zeros are replaced by the corresponding morphemes from the primary section. Examples are the sentence sequences question–answer and assertion–question (*The book disappeared # What disappeared?*). Many irregular sentence structures appear only in a context: *I can't decide where to go # Maybe to New York*. Here the second sentence is filled out to consist of the introducer *maybe*+zero variant of subject *I*+zero variant of verb *go*+P N *to New York* 'answering' the pro-P N *where*. There are also matched introducers: *Some people will come # Some won't* or *Others won't*, where the second sentence has zero for *people* and for *come*. Most conjunctions (*and*, *or*, *either . . . or*, *more than*, etc.) occur between two sentence structures that usually have a single sentence intonation, and are in many cases not recognizable as two sentences until the zeros are established. Yet another example is the *wh-* set (*The fellow whose pen you lost is back = The fellow is back + You lost the fellow's pen*).

In other combinations the second section can be shown indeed to have the constructional relations of a sentence among its parts, but the whole may not have the external form of a sentence. Zeros and pro-morphemes may not even be involved in it. Such are the V without v (but usually with *-ing* or *to*): *Being alone, he didn't go*, where zero subject *he* can be shown before *being*; *For John to win, he would have to try harder* or *To win, John would . . .* with zero *John* before *to win*; *him coming* in *I met him coming*. Such also are the nominalized sentences imbedded in others: *I resent his coming*. And such are the A N and N P N and other constructions that will be seen below to be extractable. In all these cases the sentence-like structures will be shown to be transforms of ordinary sentences; thus, *being alone* and *he being alone* are transforms of *He is alone*. [. . .]

3.3 *The question sequence*

We consider sequences of sentences in which the second has zero recurrences based on the first. A special case of this is the sequence of assertion and question (*John came here # He did?* or *Did he?* or *Who came, did you say?*), or question and answer (*Who came? # John* or *Did John come? #*

Yes, he did). In each pair, the composition of one sentence can be described in terms of its predecessor, and is indeed seen to be a transform of it. [. . .]

Since each question contains the same words as its neighboring assertion (aside from interchanging *you* and *I*), we can say that it is obtained by a transformation (from that assertion) which takes place in the presence of the question elements. In the case of the *wh*- questions, there is the additional transformation of substituting the post-*wh*- pro-morpheme as a positional variant of the constructionally corresponding section of the assertion.

3.4 Matched sentences

There are in English, in addition to the question-assertion pairs, other sequences of matched sentences, namely sequences in which both sentences contain the same words in all but one or two positions. Those sequences are characterized by having contrastive stress or reduced stress on the second sentence (the last, if there are more than two); and by having the possibility of zero variants for the recurrent words in all but one of the sentences (usually in all but the first). Since all these characteristics may be found in other successions of sentences, one could say that the ones described above are merely those successive sentences which have various words in common. However, for transformational purposes it is convenient to consider these sentence sequences separately, because they are marked by certain introducers or adverbial phrases which thus become transformation indicators. [. . .]

3.5 Conjunctions

A somewhat different type of sequence is that marked by conjunctions (*and*, *since*, etc.). Here we have two sentence structures joined by a morpheme of the conjunction class C. In the case of some members of C the two sentence structures always have a single sentence intonation extending over them: *He's taller than I am*. For many other members of C the two sentence structures sometimes have a single sentence intonation, sometimes separate intonations (often with comma intonation on the first, reduced or contrastive stress on the second): *I'll go there if I can* and *I'll go there, if I can*. Many members of C appear sometimes with matched introducers: *Some . . . But others . . .*

Another difference between conjunctions and matched introducers is that the two or more sentences joined by conjunctions may have no words in common, and yet the conjunctions or intonations may be the same as when they have. However, in those cases where the two sentences have the same words in corresponding positions, the occurrence of zeros and pro-morphemes is much the same as in matched sentences. The details as to which positions may be zeroed (or in some cases must be zeroed) differ for different members of C, and also differ when the C occurs with comma intonation. For many C the subject N is never zeroed (but the V often is): so for *because*, *since*, *if*, *as if*, *as*, *unless*, *though*, *while* (*The sailors got the raise because they had organized for it*; *The children can see the trees next fall if they don't this fall*).[15] For some C a large part of the primary sentence may be zeroed (or given in pro-morpheme or in full) in the secondary sentence; and for some C certain positions of the primary sentence are always zeroed in the secondary: *He is quick rather than careful*. For some C, if only one constituent in the secondary sentence contains different words from the primary, all the repeated words may be zeroed, and the remaining constituent may be conjoined directly to its corresponding constituent in the primary. Thus we have *He and I will come*, which we can derive from *He will come and I will come*, or *I can and will go* from *I can go and I will go*. In varying degrees this is the case for *and*, *or*, *either . . . or . . .*, *but*, *as well as*, *that is*, etc. [. . .] A complete statement for all C involves a considerable body of detail.[16] [. . .]

As in matched sentences, each reduced sentence is a transform of the complete sentence which we can reconstruct on the basis of the primary sentence, even though the reconstructed sentence has various sections which differ from the primary. Each different rule for zeroing is a different though related transformation, which takes place in the presence of the particular members of C for which that rule holds. The transformation of footnote 15 is yet another, for particular C. The original long sentence is not merely the sum of the reconstructed sentences, but the sum of these with their conjunctions.[17] [. . .]

---

[15] Some of these C occur also with a different transformation, in which the subject N is zeroed and the V has *-ing* instead of a tense or auxiliary v: *The sailors got the raise while (still) organizing for it*. When the reconstructed secondary would be N *is* N, N *is* A, or N *is* P N, this transformation may treat *is* as a V or may drop the *is*, when the subject N is zeroed: *The sailors got the raise while being on duty* or *while on duty*. In this transformation the secondary sentence ($C+S_2$) often appears first: *While walking home, we found this stone*. In *Walking home, we found this stone* we can say there is a zero member of C. With zero and certain other C, the zeroed subject N is followed by *to* V in the place of v V: *To get there we must rush*; *They rushed so as to get there on time*.

[16] Certain occurrences of *and* and *or* cannot be regarded as deriving in this way from two sentences; e.g. *Sugar and water make syrup* (compare *Sugar with water makes syrup*), *They argued back and forth*, *A black and white drawing*, *She and I disagreed*. It is generally possible to characterize these cases. Certain specific combinations of specific morphemes+*and* fall in this category, e.g. *between* N *and* N (*I lost it somewhere between Fifth and Sixth*). In the case of N *and* N V *each other* (*Electrons and positrons repel each other*) we do not directly expand the N *and* N because of the pro-N-pair *each other*.

3.7 *Constructionally included sentences*

Another type of sentence combination occurs when a transform of one sentence occupies a constructional position within another sentence: *They let the newcomer speak*, where *the newcomer speak* is object of *They let*, but neither *the newcomer* nor *speak* alone would occur as object. In the word-sharing combination the shared word fills a position in both sentences: in *I bought a house built by him*, we find *a house* both in the object range for *I bought* and also in the subject range for *built by him*. In such combinations it is not even necessary to find any grammatical relation between the two sentences, since each has its own complete structure; they simply overlap in sharing a word. In the present case, however, only the subsidiary sentence has a complete structure: *the newcomer speaks*, transformed by dropping the v into *the newcomer speak*. The primary sentence does not have a complete structure unless the subsidiary is taken as filling some status within it: *They let* does not otherwise occur without an object, and neither *the newcomer* nor *speak* but only their sum is in its object range. [. . .]

3.8 *Sentences combined by* wh-

The two types of sentences containing *wh-* (aside from questions, para. 2.62, p. 161) can be analysed as sentence sequences connected by *wh-*, with some N, P N or A of the secondary sentence S_2 replaced by a pro-morpheme which stands immediately after the *wh-* at the beginning of S_2.

In the first type, the *wh-* (plus the subsidiary sentence S_2) occurs after any N of the primary sentence S_1; the following pro-morpheme has this N (or the first N of N P N) as its antecedent, though it may have any N or P N status in its own subsidiary sentence: *The fellow who passed, The fellow for whom I got it* or *The fellow who(m) I got it for, The place in which it was* and *The place where it was*. The whole $wh\text{-}+S_2$ thus occupies a constructional position within S_1, being an A phrase of one of its N. It is comparable to the word-sharing combinations in that S_2 always has an N in common with S_1. But the transformational machinery is different, since the common word is not overlapped but repeated in pro-morpheme (something which permits the use here of more complex S_2). Other words in S_2 which are identical with words of the same status in S_1 may be zeroes as in parallel sentences. When the $wh\text{-}+S_2$ is intonationally separated it has the properties of a descriptive A: *The report, for which I'd been waiting all day,*

---

[17] *He is richer than you* is not *He is rich + you are rich* but *He is rich + -er than + you are rich*. See Sapir (1944).

finally arrived. Otherwise it can occur in matched sentences like a partitive adjective.

In this post-N type, a *wh-* word which is subject of its S_2 may have a variant *that* (*The fellow that passed*), and a *wh-* word which is object of S_2 (or part of P N with the P at the end of S_2) has variants *that* and zero (*The fellow I saw, The fellow that I got it for, The fellow I got it for*). These are subsidiary transformations of the *wh-* group. [. . .]

3.9 Summary of sentence sequences

We review all the cases where successive sentences are combined. In almost all such sequences one sentence (often the first) is not transformed and may be called the primary sentence. This can be determined for each type of combination.

There are several different co-occurrence relations between the S. Two or more sentences may each be completely independent of each other (except for transformational effects, to be mentioned below), as in some SCS. All but the primary sentence may be dependent in that some of their word co-occurrences do not appear in primaries, or in that S_1 and S_2 (with their particular word co-occurrences) do not occur in reverse order with the same connective (para. 3.2, some from para. 3.4, some SCS, first type in para. 3.8). Or all the sentences may be dependent in that each contains elements that occur only in the neighborhood of the other, as in contrasting partitive adjectives. The question sentence may be considered completely dependent, since it is identical with the assertion, except for transformations and the introductory (or connective) question element.

There are also different constructional relations between the sentences in a sequence. Two or more sentences may be grammatically independent. Or one sentence may occupy an unessential position in the other e.g. constituting an A-phrase, or a PN, or an appositional N-phrase within the other (first type in para. 3.8). Or one sentence may fill an essential position within the other, as its subject or object phrase (para. 3.7, end of para. 3.8).

There are various types of connective between the sentences. In respect to intonation there may be mere succession of two sentence intonations (para 3.2), or contrastive stress, or reduced stress or comma intonation on the secondary sentences; or the secondary sentence (in its transform) may simply be included in the intonation of the primary (e.g. *wh*-forms without comma). In respect to connecting morphemes there may be matched introducers, question markers, conjunctions between the sentences or before each one (and these 'coordinating' if the two S are reversible within the same text, 'subordinating' if not), P before a nominalized secondary sentence (footnote 15), *wh-* of both types. Finally, the overlapping of para.

3.6 may be considered a type of connection, as also the positioning of S_2 in the subject or object or other position of S_1 (para. 3.7, p. 168).

The types of transformation, mainly in secondary sentences, are as follows: pro-morphemes in all parallel sentences (para. 3.2–5, pp. 165–6), and in a special way in para. 3.8, p.168; zeroing of parallel recurrent elements in parallel sentences (with different details for different connections); inversion after the question element; and bringing the pro-morpheme constituent to the beginning of S_2 after *wh-*; dropping of v or *be* + v and in some cases adding *to* or *ing* to V; and many nominalizations of the secondary sentence.

In many cases the same S_1 S_2 sequence may occur with various connectives between them, or with various transformations on S_2 ,with the resultant meaning either virtually the same or different according to the connectives and the transformations in question.

4 The transformations of English

4.0

The following list presents the major English transformations in bare outline. Each of these accords with the definition that the same n-tuples of class members satisfy the two or more constructions which are transforms of each other. A detailed discussion of each transformation cannot be given here; the reader can test them himself, although in some cases more complicated methods may be needed in order to characterize and isolate exceptions or special cases. This is not a complete list. Perhaps it is impossible to determine a complete list for any language; but this, if it is true, will not affect either the practical or the theoretical use of transformations. The transformations will be noted in three main groupings: those that occur in independent sentences ($S \leftrightarrow S$); those that occur in sequential sentences ($S_1 \leftrightarrow S_2$); and those that occur in sentences that occupy the position of an N-phrase ($S \leftrightarrow N$). Minor groupings will be noted below.

4.1 $S \leftrightarrow S$

The main types here are passive, introducers, alternative order, and various transformations of individual words.

4.11. N_1 v V N_2 → N_2 v *be* V*en by* N_1 (passive): *The children were drinking milk, Milk was being drunk by the children.* Morphemically, *be -en* looks like one of the expanded members of v. But whereas every subject–verb pair that satisfies one member of v satisfies every other (*The children drink, The children were drinking*, etc.), they do not necessarily satisfy *be -en* (we would hardly find *Milk drank*). This *be -en* occurs after any member of v, including the expansions *be -ing, have -en*; the suffix part of all three occurs after the next verb morpheme in the sequence: *The children + ed + be -ing + drink + milk, Milk + ed + be -ing + be -en + drink + by + the children.* The N above represents any N-phrase. The V represents every V word, VP compounds (those in which the P may occur at sentence end: *They tore the paper up*), and certain other V+P: *The paper was torn up by them, His attempts were laughed at by everybody.*[18] [. . .].

4.12. *Introducers.* There are a number of individual words or word sequences which occur before any sentence, so that we have a transformation S ↔ Introducer+S. Some of these sequences are members of the classes PN or D; others are N V or N V *that*. Examples: for V which do not occur with objects, N v V ↔ *There* v V N (*A boy came, There came a boy*); and with certain exclusions, N_1 v V N_2 ↔ *There* v V N_2 N_1 (*At this point there hit the embankment a shell from our own lines*). Another such transformation is S ↔ *There* v *be*+certain nominalizations of S (para. 4.3, p. 173): *There was a barking of dogs, There was much chasing of cats by the dogs, There are the dogs chasing the cats, There will be people waiting there.*[19]

4.13. Certain similar introducers transform into an S any N (or A or P N) word or phrase (not merely one which is itself transformed from an S): *It* or *There*+v *be* (*It's a report, There will be a report*); this can be viewed as a grammatical element making N into S (an ns element by footnote 2). Combining this with other transformations (by *wh-* or word-sharing combinations) yields various types of transformed sentences: N v V ↔ *It's* N *that* V (*It's my brother that came,* also with *who* or *which* or zero in position of *that*); N v V ↔ *There's* N *that* v V; N v V+( ) *is* A ↔ *It's* A *for* N *to* V (*It's good for him to try,* via the nominalized first sentence occurring as subject of the second. *For him to try is good*); etc.

4.14. Alternative orders occur occasionally in several constructions, chiefly the following: N_1 v V N_2 ↔ N_2 N_1 v V (*The public he always despised*);

[18] But other V+P N do not have this transformation: there is no passive of *The flowers grew near the wall.* Nor does the passive occur with N *is* N sentences: *They became refugees, We are at peace.* In N *is* N sentences the long morphemes *-s* 'plural' etc. extend over both N; the position of the second N can be occupied by any N-phrase, PN-phrase, or A-phrase (including V*ing*, or V*en* from the passive); and a certain subclass of V including *become, remain,* etc. occurs in the position of *is*.

[19] There are many other introducers, with the constructional status of D, P N or N V *that*, which occur with almost any sentence: *perhaps, certainly, in general, It is probable that,* etc. We could propose transformations, for example: S ↔ *It is probable that* S; or N *v* V D ↔ *How* D N v V! *(How quickly he came!)*. Since there may be some S with which these do not occur, we have to consider them quasi-transformations; note that they contribute more specific semantic additions than do the transformations.

N_1 v V N_2 X $\leftrightarrow$ N_1 v V X N_2 (X being the rest of the nominalized sentence whose subject is N_2, or being the P of V P compounds, or certain D), as in *He threw the door open*, *He threw open the door*. When of two sentences with the same subject one is made secondary by dropping the v or v+*be* (para. 4.3, p. 173), the common subject may come before or after the remainder of the secondary: N, X, V $\leftrightarrow$ X, N V (*He, an inveterate libertarian, opposed the measure*). The two sentences here are N *is* X and N V.[20]

4.15. A few transformations can be set up in the environment of a particular subclass or individual words. For a subclass V^g which includes *give* and *tell*, $V^g N_1 N_2 \leftrightarrow V^g N_2$ *to* N_1 (*give him this, give this to him*). As an individual example: N_1 v *be more* A (or A*er*) *than* N_2 $\leftrightarrow$ N_2 v *be less* A *than* N_1 (*The sun is larger than the earth, The earth is less large than the sun*). But by no means can all logical or semantic opposites be transformationally paired in such a way. And even where this seems possible (for example between N_1 v *buy* N_2 *from* N_3 and N_3 v *sell* N_2 *to* N_1), there are usually so many special cases as to make these at best quasi-transformations.

4.2 S $\leftrightarrow$ S_2

There are quite a few transformations which change an independent sentence into a sequential one, these being the sentence combiners (or sequence markers) of para. 3.2–5 (pp. 165–7) and para. 3.8 (p. 168) with their attendant changes in the secondary sentence (including the order changes of footnote 20). This includes bound pro-morphemes in general (including *other*, many occurrences of *both*, etc.); the various distributions of zero recurrence for the various matched introducers, conjunctions, and other types of parallel sentences, including question; the specific matched introducers, question markers (with their *you/I* interchange), members of C, and *wh-*. All of these sentence-combiners have specific meanings, something which is not clearly the case for other transformations. We may consider these sentence combiners to be morphemic elements (which they generally are), and the attendant changes to be transformations. Many transformations are common to a variety of sentence-combiners, e.g. the zero recurrence. We may say then that the main transformations which are limited to S_2 position after various sentence combiners are these: bound pro-morphemes, zero recurrence, and certain order changes.

---

[20] Certain dependent order changes may be mentioned here, though they can be regarded as part of the sentence-elements with which they occur. Most of these are limited to S_2 position (after particular combining elements). The order change N v → v N is usual with the question intonation, and always with the element *wh-?*; the inversion of tenses with *not* can be expressed by assuming *not* to occur before v and adding *not* v → v *not*; and the bringing of the pro-morphemed section to the beginning of the sentence occurs always after *wh-*.

4.3 S ↔ N

The final large group contains transformations which nominalize a sentence, i.e. change to a form that can appear in one of the N-phrase positions of another sentence. Note that the transformations of para. 4.2 (p. 172) do not themselves carry any structural function, for example to combine sentences or to make them into questions; this is done by the morphemic elements that appear in each case. Here we can say similarly that the various transformations do not themselves nominalize their sentences, but that this effect is due to the appearance of the sentences in particular constructional positions within other sentences. The sentences then are N-phrases by virtue of their position; but the transformations apply only when they are in these positions.

Almost all these transformations involve dropping the v (and often the associated *be*, before A and N phrases); in most, the v is replaced by *to* or *-ing*. Particulars are given in the next thirteen subsections.

4.301. N_1 v V (N_2) ↔ N_1*'s* V*ing* ((*of*) N_2), and V*ing* ((*of*) N_2) *by* N_1 (parentheses indicate occasional occurrence): *your reading* (*of*) *these things, reading* (*of*) *these things by you.*[21]

The result has an N-phrase structure, with V*ing* as head, (*of*) N_2 and N_1*'s* or *by* N_1 as modifiers (N_1 and N_2 being themselves whole N-phrases with their own T, A, etc.); hence V*ing* may have *-s* 'plural', and if it is subject of the primary sentence this *-s* will extend to the V of the primary: *Your reading of these things is incisive, Your readings of the play are incisive.*[22] N *is* N sentences (see footnote 18) occur only in the first of these transformations, and without the *of*: *your being a writer, the door's being open.*

4.302. N v V ↔ V*ing* N, the N being the head of the resulting N-phrase.

Here the whole v, including its expansion, is dropped, so that the V is a single word. The N usually is also a single word; the T and most A that preceded N now precede the whole sequence V*ing* N: *barking dogs*; *the dangerous barking dogs.* Although this transformation differs from the preceding in various respects, homonymy between them is possible in a particular case: when nothing precedes the V*ing*, and no T precedes the N_2 above, and the N_1 above is dropped (para. 4.4, p. 175), and of course when the N happens to be in both the subject range and the object range of the V. Thus *The lobbyist visited some journalists* → *Visiting journalists* (*, he thought,*

---

[21] Here and in the other transformations it is only the v proper (auxiliary or tense) that is dropped. The expansions of v *(be -ing, have -en)* remain, and the *-ing* is added to the first of them (except that the sequence *being* V*ing* is avoided): *He has read* → *His having read, He has been reading* → *His having been reading.*

[22] After *-s* the parenthesized *of* always occurs. More restrictedly, as a quasi-transformation, we have $\acute{N}_2$-$\grave{V}$*ing* in the position of V*ing* (*of*) N_2: *your play-readings.* Another quasi-transformation is N_1 *is a* V*er of* N_2, $\acute{N}_1$ *is a* N_2-$\grave{V}$*er*, and (dropping the N_1) $\acute{N}_2$-$\grave{V}$*er*: *He is a play-reader*, etc.

might help his plans); *Journalists visited the new premier* ↔ *Visiting journalists* (, *he thought, might help his plans*). In those constructional situations in which homonymy is possible the transformation is one-directional; for, given V*ing* N, we do not know whether a pair V, N that satisfies it also satisfies N v V or also satisfies v V N. The homonymy disappears if a long morpheme appears, because of the difference in headship: *Visiting journalists was indeed helpful, Visiting journalists were indeed helpful*. The present transformation holds for V without object N. For V with object N it occurs often (as above), the object being then either dropped or prefixed to the V: $\acute{N}_2$-$\grave{V}$*ing* N_1 (often only a nonce form: *a journalist-visiting lobbyist*).

4.303. N v V ↔ V*ing of* N, with V*ing* as head: *the barking of dogs*.

Here V is as above, and homonymy is again possible with the transformation of para. 4.301 (p. 173), when its subject N_1 is dropped: in *reading of plays*, *plays* is transformed from the object (para 4.301, p. 173); in *barking of dogs*, *dogs* is transformed from the subject.

4.304. N_1 v V (N_2) ↔ N_1 V*ing* (N_2).

This does not have the internal structure of an N-phrase; it occurs as object (after V or P), but not as subject of another sentence, the N_1 having an object affix when it is a pronoun (see para. 4.308): *They found the lobbyist* (or *him*) *visiting journalists*. It also occurs after C, in which case the N_1 is subject (and often zero, if the parallel primary sentence has the same subject): *They turned off, he going first*; *Taking the hint, he went away*. From sentences of the form N *is* N we have, for example, *They turned off, he being first*.

The next group replaces v by *to* rather than *-ing*.

4.305. N_1 v V (N_2) ↔ *for* N_1 *to* V (N_2)

This transformation occurs when its sentence is subject of another, or secondary to another; despite the *for*, it is an N-phrase, not a PN-phrase: *For him to visit journalists* (*is useless*); when the transformation is from the passive, for example, we have: *He praises paintings* → *Paintings are praised by him* → *For paintings to be praised by him* (*is most unusual*). [. . .]

4.308. N_1 v V (N_2) ↔ N_1 V (N_2).

This occurs when the sentence is object of certain V (*make, let, have, feel, see, hear*, etc.) in another sentence: *He took it* → (*I let*) *him take it*. The whole v, including expansions, is dropped. Some of these V of the primary sentence also occur with other transformations of the secondary: *I saw him take it, I saw him taking it*. This transformation also occurs, with other V of the primary sentence, after *that* or its zero variant: *I insist* (*that*) *he be there* (from the sentence *He is there*), *We demand* (*that*) *he stop it*. Here the N_1 does not have the object affix when it is a pronoun. As elsewhere, the object affix is determined not by the transformation but by the relation of the secondary sentence to the primary, i.e. by the combining element, which in one case is *that* and in the other is object position; *that*+S_2 can occur as subject of S_1 as well as object of S_1.

4.309. N_1 v V (N_2) ↔ *that* N_1 v V (N_2) or with zero variant of *that*.

This occurs as subject or object of other sentences, after particular V of the other sentence: *I believe* (*that*) *he went*; *I insist* (*that*) *he is there*; *That he took it is certain.* Although the same V of the primary sentence (*insist*) may appear with this transformation and with the preceding one, it appears in the two cases as member of different V subclasses, and the meaning is different.

4.310. N_1 v V (N_2) ↔ 'N_1 v V (N_2)'.

Any sentence may appear unchanged except for the intonation of quotation, after certain V (*said*, etc.) of the primary sentence, its status being object of that V.[23]

Aside from participating in many of the v-dropping transformations above, sentences of the form N *is* N (in the sense of footnote 18) have the following transformations.

4.311. N_1 v *be* N_2 ↔ N_1 N_2 as object of another sentence (after certain V): *They considered him a police agent*; also after some members of C (*as*, zero C), in which case N_1 does not have an object affix when it is a pronoun and often is zeroed: (*As*) *a police agent, he was hunted down by the rebels.*

The transform occurs in any N position of the primary sentence when N_2 is P N, or when N_2 is marked off by comma intonation; the relation of the secondary sentence to the primary is then of the word-sharing type. We see both cases in *The fellow from Paris, a small hamlet nearby, spoke up.*

4.312. N is A ↔ A N in the word-sharing combination: *The storm is distant* → *The distant storm* (*rumbled*).

4.313. N_1 *has* N_2 ↔ N_1*'s* N_2; the result may appear (as a word-sharing combination) whenever N_2 appears.

4.4 *Many-one transformations*

Most of the preceding transformations are one-one in the sense that for each individual sentence there is one transform and conversely (except for cases of homonymy). The N_1 v V N_2 → N_2 *to* V is many-one, in that various sentences (with different subjects) have the same transform. All the v-dropping transformations may be considered many-one in this sense. Other many-one transformations follow.

---

[23] In a quoted-sentence object of N_1 *said to* N_2, *I* is a pronoun of N_1, *you* is a pronoun of N_2. The position of the quote marker can be occupied by *that* (or zero), in which case the special pronoun relations above do not hold, and certain changes occur in the v.

4.41. *Free pro-morphemes, e.g.* N_1 v V → *He* (or *She* or *It*) v V. These transformations are S → S and I(S) → I(S), where I(S) indicates transformed S. For example, *The Hungarians rebelled → They rebelled*; *The rebels' setting up councils* (*led to a shadow government*) → *Their setting up councils* . . . If I(S) is a nominalization, its pro-morpheme is a pro-S (really pro-nominalized-S), e.g. *This led to a shadow government.*

4.42. Dropping of *by* N, both in S ↔ S and in S ↔ N transformations; and the dropping of N'*s* and *for* N in the S ↔ N group. In all these cases the sections that are subject to dropping have the structural position of P N and A, positions which are sometimes filled and sometimes not filled (i.e. which may structurally be considered droppable). As a result we have by the side of *Milk was being drunk by the children* also *Milk was being drunk*; by the side of *Your reading these things* also *Reading these things*; etc. [. . .]

5 The place of transformations in linguistic structure

5.1 Elementary transformations

The difference between any two constructions which are satisfied by the same n-tuples of their word classes comprises a transformation. When we compare the constructional differences that are contributed by various transformations we may see that one such difference is the sum of two others; that is the effect of one transformation can be obtained by successive application of the other two, perhaps in a particular order. For example, N_1 v V N_2 → N_2 v *be* V*en* (*I saw him*, *He was seen*) is obtainable by the ordered succession of N_1 v V N_2 → N_2 v *be* V*en by* N_1 (para. 4.11, p. 170) and S *by* N_1 → S (dropping *by* N_1 or more generally dropping a separable P N. Furthermore, comparison of various transformations may show that these can be combined out of certain elementary changes, even if these do not occur by themselves. Any transformation which is not obtainable by combining the effect of two or more other simpler transformations will be called an elementary transformation.

5.2 The algebra of transformations

The existence of elementary transformations makes it possible to regard all transformations as compoundings of one or more elementary ones. There are a great many transformational relations among constructions, e.g. N_1 v V N_2 → *by* wh + pro-morpheme v *be* N_1 V*en ?* is a transformation (*The workers rejected the ultimatum*, *By whom was the ultimatum rejected?*). The list in 4 is relatively short because it is in approximately elementary terms; the transformation just cited is not listed, but can be obtained from para. 4.11 (p. 170) plus the question transform of para. 4.2 (p. 172).

If we now consider a transform to be the effect of perhaps several elementary transformations, rather than always of just a single transformation, we have to see in what way the various elementary transformations can occur together. Here we have to bring into consideration the fact that transformations are restricted to particular structural environments. This is most obvious for the transformations of para. 4.2 (p. 172) and para. 4.3 (p. 173), since, for example, zero recurrence appears only in secondary parallel sentences, and dropping of v without replacement by *to* or *-ing* occurs only after certain V (*make*, *let*, etc.). But it is also true for the other transformations, since, for example, the passive occurs only for sentences with N v V N structure, hence not for N *is* N.[24] In view of this, we have to recognize that a given transformation does not apply to all sentences. And a given transformation T_1 operating on an appropriate sentence S_1 may leave the structure of S_1 altered in such a way that another transformation T_2, or for that matter T_1 again, cannot be applied to the altered T_1 S_1 even if it could have been applied to S_1 itself. For example, the passive T_p carries S_1 *The Hungarian workers staged a sit-down strike* into T_p S_1 *A sit-down strike was staged by the Hungarian workers*; but to the resultant T_p S_1 we can no longer apply T_p (to yield T_p T_p S_1), because the resultant has the structure N v V P N, which is not in general subject to T_p (see footnote 18).

It is sometimes possible to analyse the observed transformations into elementary ones in more than one way; and for each different analysis, the detailed algebra of how these elementary ones are compounded to yield the observed resultants will be correspondingly different.

The successive application of elementary transformations can be called their product. For example, the sentence *May there be mentioned now a certain secret?* can be derived from N *may mention now a certain secret* by the product of transformations T_p T_d T_t T_q described in paras. 4.11, 42, 13, and 2 (question) (pp. 170, 176, 171 and 172) respectively; along the way we have T_p *S*)*A certain secret may be mentioned now by* N), T_p T_d S (*A certain secret may be mentioned now*), T_p T_d T_t S (*There may be mentioned now a certain secret*). Some products may not occur; or they may occur in one order but not in the other (e.g. T_d, dropping of *by* N, obviously occurs after but not before T_p). But this situation may be partly remedied by regarding certain transformations as positional variants of others.[25]

---

[24] In general, a given structure may be subject to several transformations, e.g. N v V N, or the object N (or any N) position in a primary sentence. And a given transformation may occur in more than one structure, e.g. zero recurrence, or N v inversion. It may be possible to characterize for a given language what types of transformations occur in what types of structures.

[25] If two transformations occur in complementary products, so that they do not both occur in the same position, we can call them positional variants of each other, and

It is clear that we have here a set of transformations with a base set (the elementary transformations), with products of the base members yielding the various other members of the set. Multiplication in the product may be associative (in the mathematical sense), and it may be commutative, depending on the properties of the transformations in the particular language and in part on how the various elementary transformations are defined. We can now define an identity transformation as one which leaves an S unchanged. An inverse can then be defined for each transformation, as that which will undo the effect of the transformation in question: the inverse T_p^{-1} of T_p is N_2 v *be* V*en by* $N_1 \rightarrow N_1$ v V N_2. The inverse T_i^{-1}, however, is of limited algebraic interest, because it usually occurs only after T_i, and will not combine with other T.

5.3 Addition (concatenation) of transformed sentences

Many of the transformations appear only when their sentence occurs with another, in some secondary status to that other. Which particular sentences occur together is a matter of co-occurrence, like which members of two classes co-occur in a sentence. But how the sentences with their sequential transformations combine is of interest as showing an operation among transformed sentences, different from that of para. 5.2 (p. 177).

We have seen three general types of sentence-combining, each with its characteristic transformations: parallel sentences, overlapping sentences (by word-sharing), and sentences nominalized as subjects or objects of others. These can be combined and recombined, so long as the conditions for applying them are satisfied; but each combination requires a new sentence to co-occur with the others. As an example we take the rather involved sentence *For the refugees from Budapest to have made so much of my arranging their border-crossing seemed to me sadder than anything else they did.* Here *The refugees cross borders* (with pronoun *they*) is nominalized as object of *I arranged* ( ). This in turn is nominalized as object of *The refugees have made so much of* ( ). This overlaps *The refugees were from Budapest*; and it also is nominalized as subject of ( ) *seemed sad to me*. The latter is conjoined by *-er than* with the zeroed parallel sentence ( ) *seemed sad to me* whose subject (the only part of the parallel sentence not zeroed) is *anything else*, to which is added by sentence overlap the *wh-* transform of *they did anything else* (with zero variant of the *which* pronoun of *anything else*, and *they* pronoun of the correspondingly placed *the refugees*). Various sentences or combinations here, including the whole, could carry various non-sequential transformations of the type of para. 5.2 (p. 176), within the limits of their application. For example, the complete sentence

---

say that they are the same transformation (but with the observed different values in the different positions).

could take the question transformation, aside from considerations of style (which could be lightened by using a different nominalization at the beginning).

5.4 The kernel

It follows from paras. 5.2 and 5.3 (pp. 176, 178) that each sentence in the language can be expressed in terms of sequential and non-sequential transformations. Given any sentence, we can check it for all transformations; we will then find the sentence to consist of a sequence of one or more underlying sentences – those which have been transformed into the shapes that we see in our sentence – with various introductory or combining elements (as in para. 3, p. 164). We have thus a factorization of each sentence into transformations and elementary underlying sentences and combiners; *the elementary sentences will be called sentences of the kernel of the grammar*. Any two different sentences will have different factorizations, either the kernel sentences or the transformations being different; but one sentence may have two different factorizations, since two one-directional transformations (applied to partly different kernel sentences) may yield the same resultant sentence (homonymy).[26]

The kernel is the set of elementary sentences and combiners, such that all sentences of the language are obtained from one or more kernel sentences (with combiners) by means of one or more transformations. Each kernel sentence is of course a particular construction of classes, with particular members of the classes co-occurring. If many different types of construction were exemplified by the various kernel sentences, the kernel would be of no great interest, especially not of any practical interest. But kernels generally contain very few constructions; and applying transformations to these few constructions suffices to yield all the many sentence constructions of the language. The kernel constructions of English seem to be only these:

N v V (for V that occur without objects)
N v V P N (for P N that have restricted co-occurrence with particular V)
N v V N
N *is* N

---

[26] The set of all sentences is closed with respect to transformations. If homonymous sentences can be suitably distinguished or marked, the set of all transformations (all the occurring products of the elementary transformations) yields a partition of the set of sentences. The set of transformations is thus a quotient set of the set of sentences. If we now map the set of sentences into the set of transformations (so that each sentence is associated with the particular product of transformations that is present in it), then those sentences which are carried into the identity transformation are the kernel of the set of sentences with respect to this mapping. These kernel sentences are precisely the underlying elementary sentences mentioned above.

N *is* A
N *is* P N
N *is* D

In addition there are a few minor constructions, such as

N *is between* N *and* N.

There are also some inert constructions which hardly enter into transformations (except quotation):

e.g. N!, *Yes!*[27]

Finally, there are the combiners and introducers of para. 3 (p. 164), and the intonationally separable introducers and D and P N phrases, all of which occur with any kernel sentence. The V in these formulas includes V P compounds and V D; the A includes D A; the N includes compound N and special forms like unitary (compound) N P N, and carries various T and post-T words (quantifiers etc.) as well as *-s* 'plural'.[28]

Different decisions at various points in the analysis of transformations will yield a somewhat different set of kernel structures. The result, however, will not make a great deal of difference for the picture of the structure of the language, and even less for the structure of language in general.

Kernel sentences contain in general no parts that repeat elements elsewhere in the sentence or in other sentences nearby; such repetition has been eliminated by the setting up of independent elements (independent within a sentence) in structural linguistics, and by the removal of dependences on other sentences in the course of transformational analysis.[29] The importance of this independence among kernel sentences is that the further grammar, the grammar of how kernel sentences are built up into the actual sentences of the language, does not have to specify very closely which kernel sentences are to be combined with which, except for the relatively loose restrictions of co-occurrence that hold among sentences. If two kernel sentences nevertheless have a special co-occurrence relation to each other; this must be stated in the grammar of constructions beyond the kernel—the grammar which contains the transformations and any other information that may be required to build all sentences.

---

[27] But V! is an N v V kernel sentence with zero variant of *you*, plus exclamatory intonation. Note that only sequential sentences, or sentences with internal evidence of zeros (as in the case of V!), are built up to full sentence structure. The traditional *Fire!* is left as it stands; we cannot know with what words to fill it out.

[28] Most cases of V*ing* will occur under V (either from *is -ing*, as an expansion of v, +V; or else from various transforms of V). But some cases are members of A; these appear in co-occurrences which are not the same as those of the corresponding V. Example: *unyielding*, and *understanding* in *He has a very understanding manner.*

[29] The sentences of a kernel are therefore maximally independent of each other; one can say that they give the least possible information about each other. (In contrast, transforms obviously contain information about each other, as do·sequential sentences.) Nevertheless, some kinds of dependences between sentences cannot be removed by the methods mentioned here, and remain in the kernel. Such are, for example, dependences among loose and unspecifiable subclasses, e.g. the partitive adjectives of a particular noun: the occurrence of one of them in a kernel sentence tells us that a partly similar sentence with another partitive adjective of that noun also exists in the kernel.

In addition to exhibiting the minimal sentence constructions, the kernel sentences are thus also the domain of the major restrictions of co-occurrence in the language. The restrictions that determine which member of a class occurs with which member of its neighbor class are contained in the list of actual sentences that satisfy each of the kernel constructions. By the nature of transformations, they do not affect these co-occurrences. The word-co-occurrences in all sentences of the language are in general those of the kernel sentences, so that the only restrictions that remain outside the kernel are the much looser ones determining which sentences combine with which. In view of this, one may raise the question whether the kernel sentences may not be subdivided further into minimal domains of co-occurrence restriction, out of which the kernel sentences would then be built according to the grammar of kernel constructions (chiefly by means of a word-sharing combination of those minimal domains). In effect, this means extracting N V, V N (object), V P N, V D and D A out of the structures which contain these (and the various compound words), and combining, say, $N_i\ V_j + V_j\ N_k$ to yield a sentence $N_i\ V_j\ N_k$.

5.5 *Effect of transformations in the language structure*

Transformations have some particular effects in the over-all structure of the language. They make possible an unbounded variety and length of sentences out of the stock of kernel sentences, thanks to the unbounded repeatability of various sequential transformations. They give an organized view of complex sentences (see paras. 2 and 3, pp. 157 and 164). As a result, they provide solutions for the structure of some constructions which are hardly solvable in the usual linguistic terms, for instance the structure of *flying planes* in *Flying planes is my hobby;* and can explain what are the differences in the two structures of a homonymous sentence (e.g. *They appointed a fascist chief of police:* both *A fascist is chief of police* transformed as object of *They appointed* ( ), and (*The*) *chief of police is a fascist* transformed as word-sharing overlap in *They appointed a chief of police*).[30]

Transformations can specify in general the differences and the similarities among sentences. Consider for example these four very similar sentences:

---

[30] Both types of problems are discussed in detail and provided with a theoretical framework in Chomsky (1955).

S_1 *Mary has a sad fate.*
S_2 *Mary's fate is sad.*
S_3 *Mary's fate is a sad one.*
S_4 *Mary's is a sad fate.*

These are transformed from some or all of three kernel sentences:

K_1 *Mary has a fate.*
K_2 *Fate is a fate.*
K_3 *Fate is sad.*

The following transformations are involved:

For S_1: K_3, overlap with K_1.
For S_2: K_1, N *has* N → N'*s* N, overlap with K_3.
For S_3: K_1, N *has* N → N'*s* N, overlap with K_2 (first N);
K_2, pro-N of second N;
K_3, overlap with K_2 (second N).
For S_4: K_1, N *has* N → N'*s* N, overlap with K_2 (first N);
K_2, zero recurrence of first N;
K_3, overlap with K_2 (second N).

It is of interest to see that these sentences, which we would intuitively describe as semantically equivalent or almost equivalent, have the same kernel sentences – except that some of them lack K_2, which we would hardly expect to contribute a semantic change – and differ only in transformations.

Transformations often overcome structural restrictions of the kernel grammar. For example, the subject of a sentence is always present in English; but it can be dropped by transforming into the passive and then carrying out the drop transformation of para. 4.42 (p. 176).[31] Or a sentence N v V, without object or even P N, can be given the passive transformation by first applying a pro-V-phrase transformation (para. 4.41, p. 176): *He stumbled unexpectedly, He did it unexpectedly, It was done by him unexpectedly.*

In many cases, transformations add flexibility in a direct way. They may change the grammatical status of a sentence into that of an N-phrase, thus making it possible to relate the sentence to an outside N V or V N, etc. They may bring out one part of the sentence for primary attention. And of course they yield stylistic variations.

5.6 Co-occurrence and transformation in structural theory

Important properties of linguistic structure are definable with respect to co-occurrence and transformations. Those constructional features of grammar which are well known from descriptive linguistics are in general limited to the kernel. In the kernel, the constructions are built up as concatenations of various included constructions, down to morpheme classes; various classes or sequences of classes (and their members) are substitutable for each other in particular positions of these constructions. Transformations cannot be viewed as a continuation of this constructional process. They are based on a new relation, which satisfies the conditions for being an equivalence relation, and which does not occur in descriptive linguistics. All sentences which are described in constructional terms must have a specific constituent analysis, since the constructional analysis proceeds in terms of immediate constituents (component subconstructions). This is not necessary, however, for all sentences in transformational analysis. Some of the cruces in descriptive linguistics have been due to the search for a constituent analysis in sentence types where this does not exist because the sentences are transformationally derived from others. For this and other reasons a language cannot be fully described in purely contructional terms, without the transform relation.[32]

---

[31] Such added flexibilities are often attained when a transformation gives the sentence the form of a different structure which has that flexibility, in this case the similarity between the passive and an N V P N sentence with omittable P N.

Some of the special operations on constructions which are set up in descriptive linguistics can be described in terms of transformations instead. For example, much of the expansion of constructions (e.g. A N = N, i.e. A N is substitutable for N) is obtainable as a result of the sentence-overlapping transformations; this will be more general if the kernel constructions are subdivided into co-occurrence domains (para. 5.4, end, p. 181). Transformations have partial similarity to certain elements and relations of descriptive linguistics. Most important of these is the similarity of transformations to variants, both free and positional. However, if A is to be a variant of B, A must be substitutable for B, or must have environments complementary to those of B, and the environments of A and B must be otherwise identical. These conditions cannot be explicitly met in the case of sentences because in descriptive linguistics environments are defined only up to sentence boundaries.

A major difference between the kernel and the rest of the language structure is that the individual co-occurrences among members of classes are in general contained within the kernel. The kinds of problems that are associated with this – for example the statistical determination of co-occurrents, the scaling of acceptance of co-occurrents, the differences among samples – all refer to the kernel. Between the kernel and the transformational structure it may be convenient to recognize a border area containing some of the types of quasi-transformations and productivity.

Finally, as has been mentioned, the kernel (including the list of combiners) is finite; all the unbounded possibilities of language are properties of the transformational operations. This is of interest because it is in general impossible to set up a reasonable grammar or description of a language that provides for its being finite. Though the sample of the language out of which the grammar is derived is of course finite, the grammar which is made to generate all the sentences of that sample will be found to generate also many other sentences, and unboundedly many sentences of unbounded length. If we were to insist on a finite language, we would have to include in our grammar several highly arbitrary and numerical conditions – saying, for example, that in a given position there are not more than three occurrences of *and* between N. Since a grammar therefore cannot help generating an unbounded language, it is desirable to have the features which yield this unboundedness separate from the rest of the grammar.

---

[32] This has been shown by Noam Chomsky (1955; 1956; 1957).

Our picture of a language, then, includes a finite number of actual kernel sentences, all cast in a small number of sentence structures built out of a few morpheme classes by means of a few constructional rules; a set of combining and introducing elements; and a set of elementary transformations, such that one or more transformations may be applied to any kernel sentence or any sequence of kernel sentences, and such that any properly transformed sentences may be added sequentially by means of the combiners. [. . .].

5.7 *Applications*

Transformations are applicable in various studies or utilizations of systems of a generally linguistic type. There are specific applications, for example, in linguistic typology – comparison of different types of language structure.[33]

The chief outside uses of transformations, however, depend upon their special meaning status. Meaning is a matter of evaluation, and cannot be fitted directly into the type of science that is developed in structural linguistics or in transformation theory. Still, for various purposes it may be possible to set up some practical evaluation of meaning; and with respect to most of these evaluations, transformations will have a special status. That many sentences which are transforms of each other have more or less the same meaning, except for different external grammatical status (different grammatical relations to surrounding sentence elements), is an immediate impression. This is not surprising, since meaning correlates closely with range of occurrence, and transformations maintain the same occurrence range. When we have transformations which are associated with a meaning change, it is usually possible to attribute the meaning change to the special morphemes (combiners, introducers, subclasses of the primary V) in whose environment the transformation occurs. To what extent, and in what sense, transformations hold meaning constant is a matter for investigation; but enough is known to make transformations a possible tool for reducing the complexity of sentences under semantically controlled conditions.

---

[33] There are also possible applications in translation, for many languages are more similar in their kernel sentences than in their total structure (the transformations, and especially the details of the transformations, being more different). Translational equivalences can be established for the kernel combinations of words, and if necessary even for the elementary transformations as such. There is also some reason to think that the kernels may function differently in memory and thought from the transformations.

References

CHOMSKY, N. (1955), *Transformational Analysis*, Unpublished University of Pennsylvania Dissertation.

CHOMSKY, N. (1956), 'Three models for the description of language', *IRE Trans. Information Theory*, vol. IT-2, no. 3, pp. 113–24.

CHOMSKY, N. (1957), *Syntactic Structures*, Mouton.

HARRIS, Z. S. (1952), 'Discourse analysis', *Language*, vol. 28, pp. 18–23.

SAPIR, E. (1944), 'Grading: a study in semantics', *Philos. Sci.*, vol. 11, pp. 93–116.

13 A. Martinet

Elements of a Functional Syntax

A. Martinet, 'Elements of a functional syntax', *Word*, vol. 16, 1960, pp. 1–10.

Whoever is concerned with the welfare of linguistic research in general and tries to transcend current parochialism must have wondered whether the label 'structural linguistics' implies anything beyond the rejection of the traditional approaches to the study of language. If we refuse to be fooled by the general use of certain terms, we may be apt to conclude that the various 'structuralistic' schools have little in common, except the attempted reduction of the phonics of the language to discrete units, the phonemes: a 'linguistic structure' cannot be the same object for a distributionalist, for whom only the respective position of the elements in the utterance is decisive, and for a paradigmatist, who concentrates on the type of relationship prevailing among commutable units. Structuralism, in linguistics, cannot be said to be a misnomer since all the structuralists operate with some sort of structure. But the use of this term by people with widely divergent outlooks certainly hampers, rather than promotes, mutual understanding and widespread cooperation.

Yet, discussions with specialists of other branches of learning and also with linguists of 'non-structuralistic' persuasion will suggest that one basic assumption is shared by those we might call the progressive linguists, namely that nothing may be called 'linguistic' that is not manifest – or manifested – one way or another in that segment of the communication circuit that lies between the mouth of the speaker and the ears of the listener. This circumstance makes contacts among members of the different structural schools worthwhile and fruitful as soon as good will is secured, whereas it is apt to prevent lasting cooperation between progressive and tradition-bound linguists.

Contrary to what seems to be felt in some quarters, this assumption does not entail that linguists should restrict their field of research to the audible part of the communication process and disregard what is not observable there: speech can only be interpreted as such, and not as so much noise, because it stands for something else that is not speech. Few linguists would maintain that an acceptable analysis of utterances could be achieved without positing that linguistically relevant variations in speech correspond to variations in meaning. Yet, the ideal of most structuralists has been to concentrate on the analysis of utterances with as few references as possible to what they stood for and, consequently, to identify and classify linguistic units on the basis of their distribution in the recorded speech segments.

This concentration on analysis represents, in my opinion, a wholesome reaction against the introspective methods favored by tradition-bound scholars. Yet, it has been detrimental when not eked out by a consideration of the communicative process as a whole. The chief aim of language is to convey information and, if – as I believe – the aim of general linguistics is to understand what language is, how a language works, how it adapts itself to new needs and, thereby, changes, no true linguistic science can disregard this fact.

An illustration may be welcome at this point: if we overlook the communicative nature of language, in other words the fact that something to be conveyed is made manifest by means of something else, we are likely to miss some basic differences between the behavior of distinctive units and that of meaningful units: if we consider two segments like *tomorrow* and *with the papers*, we will know they are distinct because one can be replaced by the other with an accompanying difference in reference, as in *he'll come tomorrow* and *he'll come with the papers*; but they may also coexist in an utterance such as: *he'll come tomorrow with the papers* or *he'll come with the papers tomorrow*. They are mutually exclusive at a certain point in the utterance, but they will be found in succession. All of this is attested in the case of phonemes: *pass* and *pat* show /s/ and /t/ to be distinct phonemes, mutually exclusive at certain points in the utterance; *past* and *pats* show that they also may coexist in succession.

There is however a basic difference between the two types of unit: using *pats* instead of *past*, i.e. the succession /t+s/ instead of the succession /s+t/ results in either changing the nature of the message or spoiling it. On the contrary, using *with the papers* after *tomorrow* or before *tomorrow* has no effect on the nature of the message. This amounts to saying that, in the case of phonemes, both their phonemic make-up and their respective position in the utterance are distinctive, whereas in the case of meaningful units or segments, the respective position of the elements in the utterance is largely irrelevant. Of course, it is relevant in many cases: *the man kills the bear* is clearly another message than *the bear kills the man*. The question the syntactician will have to answer is: How is it that what applies to *tomorrow* and *with the papers* in their relation to each other or to the rest of the utterance does not apply to *the man* and to *the bear*? In other words, why is distribution relevant here and irrelevant there?

Please note that we are not concerned here with what has been called referential meaning and carefully shunned by a majority of structuralists.

Yet, in order to formulate the problem, a central problem of syntax, we have had to keep in mind the basic communicative function of language. We have had to remember that the vocal nature of our languages imposes a linear articulation on non-linear experience: such an experience – an unpleasant experience – as a headache has nothing linear about it; but informing the doctor about it implies the linear succession *I-have-a-headache*. Speaking implies itemizing; it results from an analysis of experience into a number of elements, each corresponding to a linguistic sign. The signs corresponding to a message are, with some exceptions, ordered in a succession. The phonetic form of each sign is, in its turn, articulated into a succession of distinctive units. All the languages that have so far been studied by linguists present that double articulation, and, in my opinion, we should agree to reserve the term 'language' to such tools of communication as have it.

Reverting to the first linguistic articulation, that according to which experience is broken down into a succession of meaningful units, it is important to remember that the success of communication requires not only that it should include the requisite number of linguistic signs corresponding to the various elements of experience, but also that the relationship obtaining between the latter should find there some sort of linguistic expression: such an experience as witnessing one person introducing someone to someone else may be summarily conveyed by means of a succession of meaningful units corresponding to such elements of experience as three different persons (e.g. John, James and Peter), a happening (the introduction), a time (e.g. last night), provided the relationships between these elements be somehow indicated: information must be given, not only about the participation of, say, John in the process, but also about the nature of that participation. If the participation of John consisted in being told by James who Peter is, the nature of that participation should be indicated in English by using *to* before the formal element designating the person, say, *John*, hence *to John*.

When this type of relationship is expressed by means of a special segment of the utterance, such as *to* in the case of *to John*, we may be tempted to interpret *to* as the expression of just one more element of experience, and it is hard to see what could be objected to that interpretation. But, linguistically, the status of *to* is, as we shall see, basically different from that of *John* because it combines information with the power of conferring syntactic autonomy upon the segment that follows.

What, in language, corresponds to the relationships between the various elements of experience is what has traditionally been called 'function' when we say, for instance, that this or that word functions as a subject or an object. Function is, of course, a purely linguistic concept. In other words, function exists only in so far as it is expressed somehow in the utterance. As we know, nothing should be called 'linguistic' unless it is manifested somehow in that stretch of the communicative process that lies between the lips of the speaker and the ears of the listener. But we should not imagine that the formal expression of function will always be clearly identifiable as a given succession of phonemes. Function, as we know, is frequently marked by the respective position of the elements of the utterance. But even when its indication involves some phonemic difference, it may be impossible to locate or delimit it exactly in the utterance: in a form like Latin *homini*, we are at a loss to tell where the expression of the dative function stops and where that of the singular begins; we know that *homini* is a dative, but we also know it is a singular, because if it were still a dative, yet no longer a singular, but a plural, we would have *hominibus* with a different choice of phonemes. In such a case, the indication of the function corresponds to a meaningful unit with both meaning and form, although we may find it difficult to formulate what that form is. The difficulties that we linguists may experience in formulating something is no indication that that something does not exist and that we cannot identify it and scientifically operate with it.

Minimal meaningful units, both those indicative of function and others corresponding to the various elements of experience, have sometimes been called 'morphemes'. But since many linguists would resent applying this term to a unit whose form cannot always be pinned down, I will rather use another word, namely MONEME, which has been used with that meaning by some linguists of the Geneva school.

The study of function, as previously defined, is, in my opinion, the central problem of syntax, and the first task of the general syntactician consists in uncovering and listing all possible ways of expressing the function of a linguistic segment. There seem to be three different types, all attested in the sentence *Yesterday the President spoke in the auditorium*. The first type is to be found in *yesterday*; for simplicity's sake we shall assume that *yesterday* is a single moneme like its French, Spanish and German equivalents; this moneme corresponds to a given element of the experience which is being communicated, the one that might be defined as 'the day before this day', but it also implies that that segment of time is the one in which the related event is to be placed. In other words, *yesterday* is quite parallel to and interchangeable with a phrase like *in 1950*, where the function of *1950* is expressed by means of the specific moneme *in*; in *yesterday* we have a single moneme whose linguistically unanalysable meaning implies a given function; since *yesterday* can be placed in other positions than the initial one, after *spoke* for instance, it cannot be assumed that its function is implied by its position in the utterance. Units of this type I suggest to designate as AUTONOMOUS MONEMES.

The second type of function marking is to be found in the following stretch, namely *the President spoke*. Here the function of *the President* is the well-known subject function; the same phrase could be used with other functions, as in *I saw the President* or *I spoke with the President*, and consequently we cannot assume that *President* with or without *the* carries both its meaning and the indication of its function, which was what we found in the case of *yesterday*; *the* is no indicator of function, since replacing it by *a* as in *a President spoke* will not change the function of *President*. The only circumstance that enables us to identify it as a subject is its position before *spoke*; *the President* is thus to be labeled negatively as a NON-AUTONOMOUS PHRASE.

The third type of function marking is the most obvious one, the one we find in *in the auditorium* where *in* can be defined as a FUNCTIONAL MONEME; *in the auditorium* is an autonomous phrase with the same type of distributional latitudes as *yesterday*. Functional monemes may be 'words' such as *in*, i.e. forms which may be found separated from those whose function they mark by other forms such as *the* in *in the auditorium*, or they may be involved in inflexions, as is the case with the functional moneme 'dative' in Latin.

All this leaves us with *spoke*, which is the nucleus of a PREDICATIVE PHRASE which includes further the non-autonomous subject *the President*. The predicative phrase should be defined as what cannot be eliminated without destroying the utterance as such. There are languages where the nucleus of a predicative phrase, the PREDICATIVE MONEME, always stands in the same relation to the other monemes of the same utterance. If this predicative moneme indicates an action, the participants of the action (such as the agent, the patient, the beneficiary) will be expressed as such according to a pattern that leaves no choice to the speaker. These are the languages that do not distinguish between an active voice and a passive voice. In such cases, we may speak of a predicative function characterized by no positive mark: the predicate is the moneme by reference to which other monemes mark their functions. In a language like English, the predicative moneme may entertain different relations with the other elements of the utterance, as shown by the two utterances *the man speaks the langugae* and *the language is spoken by the man*; in which case function marking is achieved according to type three (*in the auditorium*) by means of some functional moneme. The two functional monemes are formally very complex in our present English examples. They are much simpler in their Danish equivalents: *manden taler sproget* and *sproget tales af manden*, glottal stop and *-r* in one case, *-s* in the other case.

A number of circumstances have contributed to delay the identification of functional monemes as such: first, as we have seen above when trying to analyse Latin *homini*, the formal expression of the functional moneme is frequently amalgamated with that of some other moneme, so that whoever insists on identifying linguistic units by their forms will reckon with only one unit per amalgam; the *-i* of *homini* will be considered one morpheme, which will hamper the identification of its two components on the plane of the contents. But even if this first hurdle is finally negotiated, the grammatical nature of these two monemes may induce the linguist not to separate them. By 'grammatical' I mean, of course, belonging to closed inventories, which is what we find for Latin cases and numbers.

This might suffice to confuse the issue: few linguists have become aware of the fundamental difference between functional monemes that serve to connect a segment with the rest of the utterance (*to* in *to John*, 'dative' in *homini*), whose character is, as it were, centrifugal, and other 'grammatical' elements, such as singular or plural, definite or indefinite, which help to define the value of the segment to which they are attached: *presidents*, not *president*, the plural in *hominibus*, not the singular as in *homini*; *the man*, not *a man* or *man*. The latter are MODIFIERS, with a centripetal character, whose relationship to neighboring units is of the same nature as that of a non-grammatical moneme like an adjective; a functional moneme is not a modifier, but a link.

But what has prevented even a scholar of Sapir's acumen from acknowledging the basic difference between functional monemes and grammatical modifiers is the existence, in both cases, of concord.

Concord amounts to using discontinuous and redundant forms for the expression of a certain meaning: in Latin *fortes et acerbos hostes*, the plurality of the enemies is expressed three times, and in three different places we are supposed to be informed of the nature of the participation of those enemies in the action implied by an accompanying predicate; in *discipulus uenit*, the *t* of *uenit*, which points to a subject other than speaker and interlocutor, adds nothing to what was previously implied by *discipulus*;[1] the same could be said of the *-s* of English *the man comes*. Traditionally it is assumed and often stated that concord has a connective function; it should, as it were, help the listener put together what belongs together. This view is mainly fostered by a long familiarity with classical poetry, where poets are found to have taken advantage of concord to let metrical necessities determine the respective position of words in the text, so that only concord will show, for instance, which nouns and adjectives belong together: in *incuruo terram dimouit aratro*, concord, even more than common sense, tells us that *incuruo* belongs with *aratro*. Even if it is found that

---

[1] Of course, *discipulus uenio* is not impossible; note the famous *qualis artifex pereo*. But a substantive subject is normally expected to be a third person.

incuruo aratro or *aratro incuruo* are far more normal constructions, it is a fact that concord has precedence over proximity, because no amount of syntactic ingenuity will ever connect two forms, such as *incuruum* and *aratro*, that are shown not to agree.

In the case of *discipulus uenit*, the situation is somewhat different: here the nature of the connection between *discipulus* and *uenit* would remain perfectly clear without any personal ending. Yet there are quite a few contexts of this type where concord affords the only way to a correct understanding of the message: in an utterance like *uenatores animal occidunt*, we are dependent on the *-nt* of *occidunt* for the understanding that it is the hunters who kill the beast and not the beast who kills the hunters. It is true that as soon as we speak of 'beasts' instead of a 'beast', concord is no help: *uenatores animalia occidunt* gives us no clue as to who does the killing, and we shall rely on the context or, if the context is ambiguous, fail to register the message in its entirety. In an ambiguous context, *uenatores animal occidunt* would still be understood, and we therefore must agree that *-nt* may act as an indicator of function. This does not mean however that we should consider it the formal face of one and only one 'morpheme' with a complex and best not analysed 'grammatical' function, but rather as the amalgam corresponding to different monemes; permanent and frequently redundant ones like 'third person' and 'plural', and a chance one, namely the indicator of the subject function of some plural nominal subjects such as *uenatores* whose ending fails unambiguously to mark the function. In any case, *-nt* is the signifier of several different monemes, just as are the nominal endings of Latin, and we have already seen that one and the same of these endings normally stands for both a functional moneme called case and a numeral modifier. The ablative ending of *incuruo* stands redundantly for ablative case and singular number, both being concurrently expressed by the same ending in *aratro*; but it is, by itself, an indicator of adjectival function in connection with a substantive of non-feminine gender 'in the dative singular', i.e. accompanied by a singular modifier and a 'dative' functional indicator. To people used to the expression of adjectival function by the simple syntactic trick of pre-position, the working of concord for identical purposes seems formidably intricate and hopelessly entangled with other grammatical processes. But if we concentrate on function rather than form, the mutual autonomy of the various formally amalgamated monemes will eventually stand out.

Once established, concord may be extensively used for functional purposes, but we should not imagine that it becomes established in a language in order to help listeners put together what belongs together. Whenever we have a chance to witness the appearance of new concord, we find that it does not arise out of a need to clarify the connections between the different segments of an utterance, but through the working of what is usually called least effort and what I would prefer to designate as a language economy. Language economy is, of course, prodigiously complex, and, off hand, we might believe that concord, which requires the addition of redundant elements, contradicts economy. But observation shows that, when school and prestige do not interfere, speakers faced with the choice between a shorter utterance involving adaptation to a specific situation and a longer utterance without such adaptation will normally prefer the latter: the French, who say nowadays *il ne croyait pas qu'il puisse* . . . instead of former *il ne croyait pas qu'il pût* . . . make use, in the subordinate clause, of seven successive phonemes (/kilpyis/) instead of five (/kilpy/) for the same amount of information, but they save themselves the trouble of deciding whether they should use one tense or another. Concord usually results from sticking to the same full form, whatever the context, whether it is repetitious or not, because it saves the speaker the trouble of adapting form to context: when reading was done by several people, the Romans would say *legunt* irrespective of whether the subject was specified or not. Sub-standard French *les gens ils lisent le journal* (/ . . . iliz . . . /) for *les gens lisent* . . . (/ . . . liz . . . /), which amounts to the same, is clearly a product of least effort.

All this means that, through concord, a segment may be made to stand as the expression of a functional moneme, and the very fact that we could fairly easily distinguish, in the rather complex Latin contexts we have analysed, between relational value and modifying role is a clear indication that we should keep apart what a one-sided preoccupation with forms had prevented many of our predecessors, both traditionalists and structuralists, from separating.

Some of the statements that precede might perhaps be construed as if the distinction established between functionals and modifiers was founded upon a subjective evaluation of their semantic contents. Yet this is not the case: the functional moneme may be identified as that which confers syntactic autonomy on a moneme or a phrase that is not, by itself, autonomous: the phrase *the auditorium* is not syntactically autonomous in the sense that, unless it is accompanied by a functional moneme such as *in*, *to*, or *above*, its place in the utterance will be determined by the function we want it to assume: if it is the subject function, as in *the auditorium is full*, it has to precede the predicate; if it is the object function, as in *they entered the auditorium*, it has to follow the predicate. In both cases, the phrase combines with the predicate into a predicative phrase with full syntactic autonomy: it is immaterial whether *they entered the auditorium* precedes or follows such a syntactically autonomous complement as *last night*. It is clear that using the modifier *the* or the modifier *a* before auditorium does not confer to it any syntactic autonomy. But if we add *in* before that phrase, we make it independent of its position in the utterance: we can put it in one place or another without any appreciable difference in the message: *in the auditorium, there are many seats* or *there are many seats in the auditorium.* In other words, syntactic autonomy, a distributional feature, is the criterion which enables us to distinguish between functional monemes, which are connectives, and those specifications, called here 'modifiers', which because of their grammatical status and formal comportment have, as a rule, been confused with them.

The adoption of the point of view presented so far entails important consequences, notably in typological matters. Within the present frame, it is fairly immaterial whether a functional moneme always presents the same form, as is the case with English *in*, or whether it has different allomorphs, like Italian *in*, which also appears as *n-* in *nella* for instance, or whether it is occasionally amalgamated with some other moneme, as is the case with French *à*, which combines with the masculine and plural articles into the 'portmanteau' *au*(*x*), or whether it is constantly amalgamated with some other moneme, as is found in Latin cases. What is said here of functionals also applies to other monemes, whether they be predicatives, modifiers, autonomous, or non-autonomous.

Now, a survey of the various morpho-syntactic typologies presented so far shows that even that of Sapir, which I would describe as the most elaborate and refined one, is, in the last analysis, based upon two, largely interdependent, formal features: first, the amount of allomorphic variation; second, the degree of inseparability of certain groups of units. It is not my contention that none of this should be taken into consideration in typological work. But this is, at most, but half of the story. It is all right to consider how languages manage to express various relations, but it is probably more basic to determine what those relations are, and it is highly desirable to set up language types based upon the existence or non-existence of certain relations; it is worthwhile pointing out that, in order to distinguish between active and passive, some languages make use of special inflexions, others of distinct suffixes, others still of special auxiliaries. But it is certainly more important to stress that some languages do distinguish between a passive voice and active voice while others don't.

In short, we should never assume that what differentiates one language from another is essentially and basically a different choice of the formal means of expression, but rather the type of analysis of experience it represents and the type of relationship prevailing among the corresponding linguistic articuli. In other words, whatever has to be retained of the Whorfian hypothesis applies to syntax as well as to the other aspects of linguistic structure.

14 K. Pike

A Syntactic Paradigm

K. Pike, 'A syntactic paradigm', *Language*, vol. 39, 1963, pp. 216–30.

1 Controlled redundancy in a citation paradigm

In a previous article I showed how matrix theory could be used to present some features of the relations between well-defined constructions (Pike, 1962). I claimed there that a traditional paradigm illustrates one kind of matrix on a word level but that matrix theory provides the basis for paradigms on syntactic levels as well.

Here I wish to give one small syntactic paradigm to illustrate this point. For structural data, I chose an article by James C. Dean.[1] It has the advantage of brevity, so that re-statement of its essential components can be fairly complete without being tedious. In addition, it represents a Philippine structure so different from Indo-European – or, for that matter, from American Indian languages – and so little understood by those who are not specialists in Malayo-Polynesian, that any light thrown on its structure is likely to be welcome.

Our procedure will be, first, to set up and discuss the relationships between Bilaan clause-types through a display of a crucial kernel matrix. We show derivatives of the matrix by multiplication of that kernel in various ways.

We then come to the transitive paradigm itself. This is displayed in three forms: 1. as a citation paradigm, 2. as a tagmemic-notation paradigm, and 3. as a tagmatic-notation paradigm.

The CITATION PARADIGM consists of sample clauses chosen so as to illustrate contrasts between members of the paradigm. The same morphemes are retained wherever possible. Every change in illustrative lexemes is intended to force on the reader's attention some element of structural significance. The continued use of the same lexemes allows the reader to keep a few in view – all that are needed, in principle, to illustrate the entire contrastive structure.

---

[1] Dean (1958). For convenience I have changed a few of Dean's orthographic symbols. For description of the sound system, see Dean (1955). The Bilaan dialect represented is spoken in Kablan, Tupi, Cotabato, Mindanao.

Forced memory of a large number of items and the elaborate search for substitution relations of comparable forms are deliberately avoided by this built-in redundancy of the paradigm. Simplicity of description here is assumed to include presentation designed for ease of understanding. Paradigmatic redundancy is welcomed when it serves to stimulate the reader's perception and to aid his understanding – rather than rejected by some esthetic criterion of rhetoric which proves irrelevant or harmful to the perception of linguistic pattern.

In order to help the reader to obtain insight into the structure, we wish to control all but one variable at a time, so that contrastive elements of that variable will be clearly in evidence. Lexicon, therefore, is held constant unless the structure demands lexical replacement or addition. A syntax paradigm exploits this CONTROLLED REDUNDANCY. This explains why paradigms have been used for centuries in pedagogical treatments.

Note that by these examples we are trying to illuminate and illustrate the structure. Validity of the forms,[2] as a related goal, is best assured by excerpting the illustrative material from running text. But this is equally true for all kinds of linguistic analysis. It negates neither here nor elsewhere the value of speed in reaching the tentative conclusions and directness of understanding gained from elicited paradigms.

The TAGMEMIC-NOTATION PARADIGM shows in a formal manner the internal structure of the contrastive clauses and the relations between their components. It does so in the simplest, most direct manner by an emic notation of one symbol-complex per included tagmeme – and one tagmeme per symbol. If alternative forms of a construction are possible, only one is represented here, chosen as the most frequent or as in some sense normal or basic.

The TAGMATIC-NOTATION PARADIGM, the third form used here, then supplies the details of alternative or expanded forms of the emic constructions. It gives each tagmeme of each construction variant an etic notation by symbolizing not only the functional slot involved but also the class of lexemes filling it.

Discussion of various problems in the transitive analysis will be followed by a brief presentation of a possible intransitive structure, and a few hints as to the nature of further Bilaan constructions. Finally, English will be contrasted with Bilaan at certain relevant points.

---

[2] Additional illustrations have been supplied to me by Dean when those in his article lack the redundancy of lexical items required for the paradigm. Some of these he was unable to check in the field. To be sure, therefore, the reader must use Dean's published forms, which include illustrations of the types and morpheme classes (but not always with the identical morphemes) shown here.

2 The transitive matrix

The crucial core – or kernel, or nucleus – of Bilaan structure includes a class of transitive clauses which are formally distinct and which have their formal contrasts paralleled by construction meaning.[3] By the construction differences the hearer is informed that the observer's (or speaker's) attention is oriented toward the relation between the action of the predicate and its actor, or between an action and its goal, or between an action and some other person, thing or locality relevant to it. In each of these types the physical event – the etic situation, the denotation – is constant, and signalled by the transitivity of the clause. But the emic focus – the directed attention of the observer (or speaker) to one of the relations of the activity as reported – becomes contrastive. Some one of the substantive components of the clause serves as the FOCUS-COMPLEMENT of this FOCUSED ACTIVITY-RELATION of the predicate and often is formally marked as such.[4]

In an English translation of a Bilaan clause one may add to the predicate a subscript $_{\text{af}}$ for an actor-focused activity relation, but $_{\text{gf}}$ for a goal-focused activity relation and $_{\text{rf}}$ for a referent-focused relation and an arrow pointing toward the actor or goal or referent which is the complement of that activity relationship. The contrastive translations then may appear as: *(It is the) cat* ← *(that) catches*$_{\text{af}}$ *(the) rat*. *(It is the) rat* ← *(that the) cat catches*$_{\text{gf}}$. *(It is for the) man* ← *(that the) cat catches*$_{\text{rf}}$ *(the) rat*.

The first two clauses might also be translated as English active and passive: *(The) cat* ← *catches*$_{\text{af}}$ *(the) rat*. *(The) rat* ← *is caught*$_{\text{gf}}$ *(by the) cat*. But deep-seated differences in structure between an English passive clause and a Bilaan goal-focused-activity clause make the former translation for the moment preferable. (In section 7 I shall outline some of these important structural differences. First, however, I wish to present the formal structure of the Bilaan clause.)

We are now ready to set up a simple column matrix listing these three formal clause types, named for the contrastive forms of their activity relationships. In (a) I use M_k to represent the kernel (the nucleus) of the activity-clause system. TC represents transitive clause (with actor-focused, goal-focused, or referent-focused activity). Our later formulas justify the treatment of the first two of these constructions as well-defined units[5].

---

[3] Sometimes called voice. See Howard McKaughan (1958).

[4] The term FOCUS applied directly to the actor or to the goal is found in A. Healey (1958) and in Phyllis M. Healey (1960). The latter is based on a master's thesis written at the University of Sydney, 1958.

Some students of Philippine language – among them McKaughan and Wolfenden – tell me that this use of focus makes the actor or goal too prominent. This in turn leads to difficulty in distinguishing between focus and emphasis, and fails to give central attention to the activity itself and its relationships. For this reason I have developed the phrases 'focused activity-relation' and 'focus complement' in order to place priority on the verb.

$$M_k = \begin{array}{ll} TC_{af} & (a) \\ TC_{gf} & \\ {}^*TC_{rf} & \end{array}$$

The third construction is marked with an asterisk because Dean's material does not provide the data necessary to treat it here. In Abrams' data, however, a small residue in the system contrasts TC_{gf} with TC_{rf}. The referent type is not needed for our presentation of a model for a syntax paradigm, and will not be mentioned again until later in the paper; but it must be kept in mind by those interested in comparing Philippine systems with English.[6]

In addition to structural attention focused on activity relations, Bilaan gives formal attention to emphasis. Focus is not emphasis. Focus reports the observer's attention to one of several relations – without essential emotional overtones – between a predicate and some other part of a clause; the focus-complement substantive is viewed only in reference to that relationship, not as in focus of itself. In emphasis, on the contrary, some one substantive is singled out for a direct isolated overlay of emotional connotation without formal (emic) reference to or dependence upon its relation to the activity to which it is in (etic) fact related. The formal independence of emphasis allows it to function as a variable which is formally separate from the focus complement.

In English, emphasis is usually achieved by suprasegmental intensity, length, and pitch (although it can also be handled in part by special segmental lexemes such as *indeed*). In Bilaan the emphasis – like a focused activity-relation – is signalled segmentally. We translate the Bilaan emphasis into English orthography, therefore, by small capital letters representing any unspecified variety of English suprasegmental intensity: *(It is the) cat* ← *(that) catches*$_{af}$ *(the)* *RAT*. *(It is the)* *CAT* ← *(that) catches*$_{af}$ *(the) rat*. *(It is the) rat* ← *(that the)* *CAT* *catches*$_{gf}$. *(It is the)* *RAT* ← *(that the) cat catches*.

---

[5] Any well-defined unit must be seen in reference (i) to the contrastive components which set it apart from other units and which simultaneously allow its identification; (ii) to the variant forms in which it physically occurs; and (iii) to its distribution in a class of units, in a hierarchical sequence (or string or spacial display) of units, and in a matrix (or contrastive pattern) of units. Thus the units in (a) comprise a matrix and are in part defined, in turn, by their place in this matrix pattern. See also Pike (1962).

[6] Norman D. Abrams (1960). This contains many details of the general verb morphology and of the morphological and syntactic aspects of focus. It amplifies Dean's shorter paper, and implies types of verb bases not treated there. The dialect described is that of the Sanangani district of southeast Cotobato, and of the Koronadal area.

In the first two illustrations, *cat* as focus-complement actor is first non-emphatic, then emphatic. Similarly *rat*, as focus-complement goal in the second pair of illustrations, varies in emphasis. We see, therefore, that the marked activity-focus relation is essential to the basic contrasts of the kernel matrix, whereas emphasis is an optional component which may be added to any of the substantive tagmemes of the clause without affecting its kernel structure.

For this reason it proves useful to treat emphasis as adding a class of further but less basic clause types to the system. This is shown in (b) as matrix multiplication, with $_e$ representing emphasis: $_{ea}$ emphatic actor, $_{eg}$ emphatic goal.

$$_e \cdot M_k = {}_{ea}TC_{af}, {}_{ea}TC_{gf}, {}_{eg}TC_{af}, {}_{eg}TC_{gf} \quad \text{(b)}$$

If, now, we arrange (b) in matrix form $_eM$, we obtain (c); we arbitrarily retain non-emphasis in the matrix by adding a zero-emphasis column. Here A_{fc} is actor-focus complement, G_{fc} goal-focus complement.

Transitive Clause $_eM$: (c)

$$\begin{array}{l|lll} & _e\text{Zero} & _eA & _eG \\ \hline A_{fc} & TC_{af} & {}_{ea}TC_{af} & {}_{eg}TC_{af} \\ G_{fc} & TC_{gf} & {}_{ea}TC_{gf} & {}_{eg}TC_{gf} \end{array}$$

But (c) does not include some apparently intransitive forms of Dean's data. These have actor-focused activity forms without the complexity of a substantive goal elsewhere in the clause. Why not start with the simpler intransitive clauses as the kernel matrix rather than the transitive? The answer seems to be that grasping the pattern of Bilaan or many other Philippine languages requires an appreciation of focus contrasts as basic. This cannot be demonstrated unless one begins with the transitive, where both actor-focused activity and goal-focused activity come into view.

This judgement then forces the question: can the intransitive matrix be derived from the transitive? It can indeed. But to do so it is necessary to multiply the transitive matrix by a component which applies a minus feature for goal. If $-TC_{gf}$ is used to mean the obligatory absence of the goal-focus cell of M_k, and Int to mean intransitive, then the non-emphatic and the emphatic intransitive clauses are seen in (d) and (e).

$$\text{IntC} = \text{Int} \cdot (M_k - TC_{gf}) \quad \text{(d)}$$
$$_e\text{IntC} = {}_e \cdot (\text{Int} \cdot [M_k - TC_{gf}]) \quad \text{(e)}$$

That is, the intransitive clause matrix is derived from (c) by changing T to Int and by striking out the bottom row and the righthand column.

3 The transitive citation paradigm

We are now ready to present in (f) the transitive syntactic paradigm in citation form. The citation forms are kept as uniform as possible, for reasons given in section I. Words illustrating the same tagmeme or tagmeme class or tagmeme variant (allotagma) are wherever possible kept in a vertical column. Gaps in the permitted occurrence of the tagmeme or allotagma are shown by dots.

The rows in (f) come in pairs, representing the basic contrast of actor-focused versus goal-focused activity. Symbols at the left in (f) come from matrix (c). The presentation of pairs is chosen successively from each column of (c).

The *M* (or *am*) in the verb *kamfe?* 'catch' is part of a variant morpheme which helps to signal the actor-focused activity construction; the comparable goal-focused activity in the same verb is seen signalled by *N*. As in the English sentences of section II, an arrow pointed toward a Bilaan pronoun shows it to be a focus complement; $_{\mathrm{af}}$ and $_{\mathrm{gf}}$ label activity-focus types; capital letters show emphasis.

Sentences (f5, f5′), with asterisk, show how the analogous pattern might possibly be filled – but without documentation in available data.

Various observations can be made by scanning the paradigm and its translation. Note that in the first permitted structural position only emphatic forms occur – but that an emphatic goal optionally comes last (f6, f7, f8). The emphatics either are out of focus or serve as the focus complement. The third singular emphatic pronominal form, actor or goal, is *kanen* 'he, him', whether or not it serves as focus complement; third plural emphatic is *dale* 'they, them'. Predicates, in the second position, contrast by their markers for actor-focused versus goal-focused activity.

More difficult to see, but present in the data, is the crucial membership in four pronoun classes, and the priority ranking of their functions.[7] Some such characteristic seems to be typical of many Philippine languages. Highest structural priority in descriptive order of selection goes to the emphatics (*kanen*, *dale*). Regardless of focus or of actor-versus-goal function, the emphatic set is used when two or three of the functions are simultaneously present. Second priority goes to the focus-complement set represented here by # 'he, him' and *ale* 'they, them'. Whether the pronoun is in actor or goal function this set is used if it carries simultaneous focus-complement function. Zero priority goes to two contrastive sets of pronouns when they are neither emphatic nor serving as focus complement.

---

[7] The emphatic set, however, is ambiguous with the non-focus object set except for third person singular *kanen* versus *dun*. Otherwise the contrast is neutralized. (For complete lists, see Dean, 1958.) Dean feels, however, that the contrast is vital to an understanding of the system.

Transitive clause in citation paradigm

(f1)	TC_{af}	...	*kaMfe?*	*→ale*	*dun*	'...	Catch$_{af}$	→they	him'
(f1′)		(...	*kaMfe?*	*→#*	*dale*	'...	Catch$_{af}$	→he	them')
(f2)	TC_{gf}	...	*kaNfe?*	*-la*	*→#*	'...	Catch$_{gf}$	they	→him'
(f2′)		(...	*kaNfe?*	*-an*	*→ale*	'...	Catch$_{gf}$	he	→them')
(f3)	${}_{ea}TC_{af}$	*DALE←*	*kaMfe?*	...	*dun*	'THEY←	catch$_{af}$	...	him'
(f3′)		*(KANEN←*	*kaMfe?*	...	*dale*	'HE←	catch$_{af}$	...	them')
(f4)	${}_{ea}TC_{gf}$	*DALE*	*kaNfe?*	*-la*	*→#*	'THEY	catch$_{gf}$	they	→him'
(f4′)		*(KANEN*	*kaNfe?*	*-an*	*→ale*	'HE	catch$_{gf}$	he	→them')
(f5)	${}_{eg}TC_{af}$	**KANEN*	*kaMfe?*	*→ale*	...	'HIM	catch$_{af}$	→they	...'
(f5′)		*(DALE	*kaMfe?*	*→#*	...	'THEM	catch$_{af}$	→he	...')
(f6)	~	...	*kaMfe?*	*→ale*	*KANEN*	'...	catch$_{af}$	→they	HIM'
(f6′)		(...	*kaMfe?*	*→#*	*DALE*	'...	catch$_{af}$	→he	THEM')
(f7)	${}_{eg}TC_{gf}$	*KANEN←*	*kaNfe?*	*-la*	...	'HIM←	catch$_{gf}$	they	...'
(f7′)		*(DALE←*	*kaNfe?*	*-an*	...	'THEM←	catch$_{gf}$	he	...')
(f8)	~	...	*kaNfe?*	*-la*	*→KANEN*	'...	catch$_{gf}$	they	→HIM'
(f8′)		(...	*kaNfe?*	*-an*	*→DALE*	'...	catch$_{gf}$	he	→THEM')

In zero-priority the actor function (a suffix class) is expressed by *-an* 'he', *-la* 'they'; the goal function by *dun* 'him', *dale* 'them'.

Note, then, that only *kanen* and *dale* stand in first position, with or without an arrow; only *-an* and *-la* in third, never with an arrow; # and *ale* both in third and in fourth, always with an arrow. This specialization of pronouns for tagmemic function characterizes Bilaan.

Such an arrangement of pronominal sets, however, raises theoretical questions. What is the morphemic status of these forms? What is their tagmemic status? And how must # be analysed?

The ranking of pronouns into functional sets affects all the persons. Yet no one person shows overt contrasting forms for all four sets. Third person singular has four forms in contrast, but one of these (the focus-complement third person) is zero. Nevertheless, by using both third singular and third plural in the citation paradigm, the four sets can be filled out overtly and contrastively for at least one of the persons for each function.

In this presentation we are assuming that the third-person-singular morphs of each functional set are contrastive morphemes, not allomorphs of a single morpheme. This solution may appear surprising, in view of the fact that elsewhere I have treated English *he* and *him* as allomorphs of {*he*} tagmemically conditioned (Pike, 1954; 1960). The Bilaan analysis is forced on us 1. because of the formal correlation of pronoun focus-complement with marked verb structure; 2. because of the form-meaning contrast of focus-complement versus non-focus-complement pronouns in identical tagmemic slots, as in (f1) versus (f2); and 3. because structural ambiguities would result if all the forms of one person were rewritten with a single morphophonemic spelling. In such a rewrite the contrastive pronominal forms (f1, f2, f8) would be the same, concealing overt formal differences which in fact parallel the semantic ones.

This conclusion, in turn, leads to # 'he←, him←' as having some formal status. We do not treat it as a zero morpheme, for no overt allomorph would then be part of it. (Tagmemic theory rejects all totally zero units.) Our solution, therefore, is to assume that the tagmeme of actor here has *ale* as an overt manifesting variant, and # as a covert variant. Then # becomes a symbol, directly, of the presence of the actor tagmeme in zero-variant form. Thus # in (f) is not a zero morpheme, but a zero allotagma. The dots, on the other hand, represent neither a zero allomorph nor a zero allotagma but merely the absence of any manifestation of a tagmeme at a given point, and the absence of any compelling reason to expect a tagmeme there.

All of this, however, leads to a still deeper structural issue. If # and *ale* in (f1, f5, f6) manifest an actor tagmeme, what tagmeme do *-la* and *-an* manifest in (f2, f4, f7, f8)? Do both pronoun sets represent precisely the same tagmeme of actor? Are they not rather functionally – tagmemically – somehow in contrast? And is not the actor manifested in (f4) by *kanen* and *dale* in contrast with goal tagmeme manifested in (5) by *kanen* and *dale* in the same apparent position? And are not all of these, in turn, in some kind of functional contrast with *KANEN* and *DALE* as emphatic actor and emphatic goal, as in (f3, f7)? Finally, what should be done to formalize the fact that # and *ale* sometimes manifest actor (f1, *f5, f6) and sometimes goal (f2, f4)?

My solution to these problems is to set up two special tagmemes, neither of which ever has a manifestation apart from another tagmeme. Emphatic tagmeme is characterized 1. by prepredicate position (f3, f4) with free distribution-variant under goal forms to post-actor position (f6, f8); 2. by the special set of emphatic pronouns, *kanen*, *dale* as filler of its slot; 3. by a semantic component of emphasis; 4. by simultaneous (portmanteau) manifestation with actor (f3, f4), or simultaneous goal (*f5, f6, f7, f8). Tagmemes in Bilaan, then, can be simultaneously present in tagmeme composites, as in other languages we may find simultaneous phonemes or simultaneous morphemes.[8]

A similar analysis must then be adopted so as to set up a tagmeme of focus-complement. This one can be simultaneous with actor or with goal tagmeme, or with the tagmeme composite actor-plus-emphasis or goal-plus-emphasis. The focus-complement tagmeme is characterized 1. by the manifesting pronoun class #, *ale* (or by overranking emphatic *KANEN* or *DALE*); by postpredicate actor or goal positions (or by overranking prepredicate emphatic position); 3. by correlation with a focus indicator for actor or goal in the predicate; and 4. by a semantic component of focus complement.[9]

The actor tagmeme outside a tagmeme composite is seen in (f2, f4, f7, f8). Characterizing the simple actor is 1. the non-emphatic, non-focus-complement pronoun class *-an*, *-la*; 2. position immediately after the predicate; 3. a semantic component as actor; and 4. negative concord (i.e. no concord) with goal-focused predicate tagmeme. (But as part of an emphatic tagmeme composite, or of a focused-emphatic tagmeme composite, it takes on the priority characteristics referred to under emphasis and focus.)

Comparably, the simple non-composite goal tagmeme of (f1) and (f3) has

---

[8] See Velma B. Pickett (1960). She sets up for various purposes simultaneous tagmemes in Double Function. See pp. 69–71, 78–9, for relative or interrogative tagmeme simultaneous with subject, object, time, location and manner tagmeme. See pp. 56 and 35 for her clause and phrase formulas in tagmemic notation. Presumably English *who* etc. could well be handled by related techniques. See also Pike (1960).

[9] Or of 'topic' – the term used by McKaughan (1958, p. 18) for the meaning and name of the focus complement. See footnote 4.

1. a manifesting class of non-emphatic non-focus-complement pronouns *dun*, *dale*; 2. position after non-emphatic actor; 3. a semantic component as goal; and 4. negative concord with an actor-focused predicate.

How can all this extra information by symbolized? The citation paradigm appears deceptively simple. Some kind of notation is needed to show the simultaneity of tagmeme structurings with its intricate priority interplay of pronominal sets. For this we turn to a tagmemic notation.

4 The structure paradigm

If one modification is made, the simultaneity of tagmemes in the clause string can be shown in (g) by symbols introduced in earlier sections. We add a vertical bar between letter symbols of a tagmeme composite, thus $A|_{fc}$, as representing simultaneous actor and focus-complement tagmemes.

Transitive clause in tagmemic-notation paradigm (g)

(g-f1)	TC_{af}	$\{\ldots$	TP_{af}	$A\vert_{fc}$	$G\}$
(g-f3)	$_{ea}TC_{af}$	$\{_{e}\vert A\vert_{fc}$	TP_{af}	$\ldots$	$G\}$
(g-f5)	$_{eg}TC_{af}$	$\{_{e}\vert G$	TP_{af}	$A\vert_{fc}$	$\ldots\}$
(g-f2)	TC_{gf}	$\{\ldots$	TP_{gf}	A	$G\vert_{fc}\}$
(g-f4)	$_{ea}TC_{gf}$	$\{_{e}\vert A$	TP_{gf}	A_{cr}	$G\vert_{fc}\}$
(g-f7)	$_{eg}TC_{gf}$	$\{_{e}\vert G\vert_{fc}$	TP_{gf}	A	$\ldots\}$

I have rearranged the paradigm, dropping the grouping into pairs of contrasting focused activity gained from the columns of matrix (c). I have regrouped the clause types into two triples by way of the two rows of (c), so as to bring out further relationships between the actor-focused set as a whole and the goal-focused set as a whole. Cross reference of the formulas of (g) to those of (f) is provided by using a composite number, (g-f4) labels the new treatment of (f4).

This notation points up a further fact: that the emphasis tagmeme has a specific position before predicate but the focus-complement tagmeme has none. Moreover, the lack of a vertical bar within the complex symbol for predicate shows that we consider the actor-focused predicate to be a single tagmeme (on the clause level of structure) in contrast with the goal-focused tagmeme. Within the predicate, on the other hand, there is a word-level string of tagmemes. In a full-scale grammar we would here have to discuss 1. the tagmeme slots within each kind of verb structure, plotting each kind of verb construction for this purpose separately; 2. the morpheme classes entering the slots; 3. the contrasts between various kinds of verb constructions; 4. relations between the verb constructions (or their included constituents) and noun (or other) structures; and 5. any allomorphic variations within morpheme classes or constructions.

All of this is morphological discussion and lies outside the scope of the present paper (and of Dean's) except for a general reference to the essential -M- and -N- signals of focus. Here we have ignored all morphological variants of these signals. Our concern with internal components of the verb has been exclusively with samples of those units which signal the interlocking of the morphological with the syntactic level — i.e. of word with clause. (If we were analysing the verb in detail, we would expect to set up a tagmeme of focus, on the morphological level. The morphemes represented by -M- and -N- would fill this tagmemic slot.) For a description of the elaborate morphophonemics of a Philippine language note rather McKaughan's treatment (1958); for our purpose the morphological detail would at this point only obscure the syntactic relations.

The overall simplicity of the formulas of (g) is due to their emic representation. Variants in the position of tagmemes, variants in the internal structure of the tagmemes (the etic components of the tagmemes in terms of filler classes or their variants), clause variants which result from clause expansions through optional tagmemes of time, location, and so on (not discussed by Dean) – all of these complicating characteristics have been omitted. The fact that we are listing stripped-down emic summaries of constructions through formulas without tagmeme variants and without construction variants is symbolized in (g) by the braces enclosing each emic clause formula.

One symbol of (g-f4), so far unexplained, is also important to the structure. The notation A_{cr} represents a special cross-referent tagmeme filling the regular actor slot when tagmeme A occurs in ${}_{e}|A$ of ${}_{ea}TC_{gf}$. Here two actor tagmemes are required. The second actor tagmeme occurs in cross reference to the first. The permitted classes in the two actor slots are not completely interchangeable. The postpredicate actor tagmeme is the more restricted of the two. Only the non-focus-complement actor pronoun may come there. But either a noun (or noun phrase) or a pronoun will be found manifesting A (with or without ${}_{e}$ or ${}_{fc}$). We are forced to give attention to the particular classes which are allowed to enter the various slots.

The emic listing of clause constituents in (g), furthermore, does not provide symbols for the variants of structure which arise from freedom in the placement of the tagmeme composites ${}_{e}|G$ and ${}_{e}|G|_{fc}$ from (f5 $\sim$ f6, f7 $\sim$ f8). These composites occur either in the prepredicate slot normal to the emphasis tagmeme or in the postactor slot normal to nonemphatic goal. (An alternate analysis might deny the presence of $|_{fc}$ in ${}_{e}|A|_{fc}$ and in ${}_{e}|G|_{fc}$ since the only signal for its presence comes from the verb. Dean, however, marks such items as still in focus; I follow his analysis.)

Matrix (h) now supplies this lack of symbols for etic detail in (g) with a tagmatic slot-and-class paradigm. The fact that it is an etic display of the constructions, rather than an emic one like (g), is symbolized by enclosing each etic clause formula in square brackets. Parentheses enclose the fillers of the slots. In predicate slots one finds a transitive verb with the actor-focus or goal-focus markers *-m-* or *-n-*. In all other slots except A_{cr} one finds both noun and pronoun.

The labels Pr_e, Pr_{fc}, $Pr_{uf}A$, $Pr_{uf}G$ denote emphatic pronoun, focus-complement pronoun, non-focus-complement actor, and non-focus-complement goal pronoun. Pointed brackets enclose one typical member of the class to remind the reader of the make-up of the class, and as a key to Dean's description, which labels the classes by these particular examples.

Paradigm (h) may also be called an etic TESTING MATRIX, since with this kind of data (supplemented with comparable matrices for word constructions, phrase constructions, and construction of sentence, paragraph, and discourse at their respective levels) the accuracy of a syntactic description and the productive applicability of the system in speech can be matched for fit against the material available for analysis.[10]

Yet (h) calls for further particulars to demonstrate its full validity. It predicts the occurrence of nouns in all A (except A_{cr}) or G positions – whether or not in simultaneous occurrence with ${}_{e}|$ and $|_{fc}$. It predicts also some ambiguity when nouns are used. Inasmuch as the symbol N remains unchanged in tagmeme slots A, G, $A|_{fc}$, $G|_{fc}$, ${}_{e}|A$, ${}_{e}|G$, ${}_{e}|A|_{fc}$ and ${}_{e}|G|_{fc}$, it implies that the nouns (or noun phrases) in those slots are identical. If this is true, then when the nouns *kukuʔ* 'cat' and *ungeh* 'rat' are used, the forms (i-f3) and (i-f5) become identical. One cannot tell here whether it is actor or goal which begins the clause or follows the verb.

(i-f3) ${}_{ea}TC_{af}$: *kukuʔ kamfeʔ ungeh* '(The) CAT ← catches$_{af}$ (the) rat.'
(i-f5) ${}_{eg}TC_{af}$: *kukuʔ kamfeʔ ungeh* '(It is the) CAT (that the) rat ← catches$_{af}$.'

This ambiguity does not occur in goal-focused clauses, however, since ${}_{ea}TC_{gf}$ (h-f4) is distinguished from ${}_{eg}TC_{gf}$ (h-f7) by the obligatory presence of a crossreferent actor pronoun in the former.

Further ambiguity develops at another place in the paradigm, but does not appear directly through (h). Its source lies below the emic level of the clause construction and even below the level of tagmemic representation

---

[10] The same data must be known, I believe, tacitly or explicitly, for an adequate test, regardless of the theory under which the analysis is made or the style in which it is presented.

Note that this claims no more than when we affirm that a phonemic description cannot be checked for accuracy against the data, nor used for producing accurate pronunciation, unless the (emic) phonological symbols are accompanied by an (etic) commentary or by an (etic) symbolism relating the systemic generalizations to physical articulation. No variety of phonemic theory can avoid this requirement.

Transitive clause in tagmatic-notation paradigm (h)

(h-f1)	TC_{af}	[…	TP_{af} (TV_m)	$A\vert_{fc}$ (N/Pr_{fc} ⟨*ale*⟩)	G ($N/Pr_{uf}G$ ⟨*dun*⟩)]
(h-f3)	${}_{ea}TC_{af}$	$[_e\vert A\vert_{fc}$ (N/Pr_e ⟨*kanen*⟩)	TP_{af} (TV_m)	…	G ($N/Pr_{uf}G$ ⟨*dun*⟩)]
(h-f5)	${}_{eg}TC_{af}$	$[^*{}_e\vert G$ (N/Pr_e ⟨*kanen*⟩)	TP_{af} (TV_m)	$A\vert_{fc}$ (N/Pr_{fc} ⟨*ale*⟩)	…]
(h-f6)	~	[…	TP_{af} (TV_m)	$A\vert_{fc}$ (N/Pr ⟨*ale*⟩)	${}_e\vert G$ (N/Pr_e ⟨*kanen*⟩)]
(h-f2)	TC_{gf}	[…	TP_{gf} (TV_n)	A ($N/Pr_{uf}A$ ⟨*-an*⟩)	$G\vert_{fc}$ (N/Pr_{fc} ⟨*ale*⟩)]
(h-f4)	${}_{ea}TC_{gf}$	$[_e\vert A$ (N/Pr_e ⟨*kanen*⟩)	TP_{gf} (TV_n)	A_{cr} ($Pr_{uf}A$ (*-an*⟩)	$G\vert_{fc}$ (N/Pr_{fc} ⟨*ale*⟩)]
(h-f7)	${}_{eg}TC_{gf}$	$[_e\vert G$ (N/Pr_e ⟨*kanen*⟩)	TP_{gf} (TV_n)	A ($N/Pr_{uf}A$ ⟨*-an*⟩)	…]
(h-f8)	~	[…	TP_{gf} (TV_n)	A ($N/Pr_{uf}A$ ⟨*-an*⟩)	$G\vert_{fc}$ (N/Pr_e ⟨*kanen*⟩)]

of etic construction variants of the clause. It arises in an etic variant (an allotagma) of the tagmeme $|_{fc}$ simultaneous with A or G (but never with $_{e}|$).

To solve this problem we have available, above, the relevant data and the requisite theory. We combine forms (f1′, f2, f5′, f6′) from the citation matrix with a discussion of third-person allotagmatic # of $A|_{fc}$ and of $G|_{fc}$ of (g) and with a discussion of the nominal ambiguity of noun function in (i-f3, i-f5).

The emic symbol for the fillers of a tagmemic slot does not differentiate between the overt and covert allotagmatic fillers of that slot. Similarly, the label Pr_{fc}, standing for the emic class of focus-complement pronouns, does not symbolize the overt and covert etic variants of the class. And finally, the selected illustrative member of that emic class ⟨*ale*⟩ does not display the zero. Where, then, would the zero allotagma (zero manifestation) of Pr_{fc} appear in the formulaic description? Where, that is, except in the citation paradigm (f)?

Tagmemic theory is explicit on this point. Tagmemes, as well-defined units, include in their complete description any needed reference to their variants (as well as to their contrastive-identificational features and to their distribution in class, in sequence, and in matrix).[11] Just as we have used braces and brackets to distinguish the emic from the etic nature of construction formulas so now we use the same symbols for the tagmemes themselves. Braces enclose an emic notation of tagmemes; brackets enclose symbols for allotagmas. This notation becomes important whenever the analyst is discussing the tagmemes as such rather than the tagmemically represented construction variants. The tagmeme composite $\{A|_{fc}\}$ for example has among its variants both $[A|_{fc}$ ⟨*ale*⟩] and $[A|_{fc}$ ⟨#⟩]. (We could have further differentiated matrix (g) from (h) by italicizing the tagmeme symbols of (g) but not the same symbols when representing only the tagmemic slots of (h). We have assumed however that braces and brackets make this difference explicit enough just as we use only one set of symbols for phonemes between diagonals and for phones between brackets).

5 Intransitive paradigm

Now that our basic approach has been illustrated and paradigmatic symbols have been set up for handling the transitive clause the intransitive clause can be presented without much comment. The intransitive citation paradigm appears as (j), p. 209. A new set of stems appears in (j) illustrated by *mngel* 'to cry'. It is the correlation of this stem type with the absence of goal forms which forces the symbolizing of intransitive actor-

---

[11] See footnote 5.

Intransitive clause in citation paradigm (j)

(j1)		IntC$_{af}$	...	*Mngel*	→*ale*	...	'...	Cry$_{af}$	→they	...'
	(j1′)		(...	*Mngel*	→#	...	'...	Cry$_{af}$	→he	...')
	(j1″)		(...	*Mngel*	→*kuku?*	...	'...	Cry$_{af}$	→cat	...')
(j2)		$_{ea}$IntC$_{af}$	*DALE*	*Mngel*	→(...)	...	'THEY	cry$_{af}$	→(...)	...'
(j3)			*DALE*	*Mngel*	→*ale*	...	'THEY	cry$_{af}$	→they	...'
	(j2′-3′)		(*KANEN*	*Mngel*	...#	...	'HE	cry$_{af}$	→#	...')
	(j2″-3″)		(*KUKU?*	*Mngel*	→#	...	'CAT	cry$_{af}$	→#	...')
(j4)		~	...	*Mngel*	→*ale*	*DALE*	'...	Cry$_{af}$	→they	THEY'
	(j4′)		(...	*Mngel*	→#	*KANEN*	'...	Cry$_{af}$	→he	HE')
	(j4″)		(...	*Mngel*	→#	*KUKU?*	'...	Cry$_{af}$	→it	CAT')

Intransitive clause in tagmatic-notation paradigm (l)

(l-j1)	IntC$_{af}$	[...	+IntP$_{af}$ (IntV$_m$)	+A|$_{fc}$ (N/Pr$_{fc}$ ⟨*ale*⟩)	...]
(l-j2–3)	$_{ea}$IntC$_{af}$	[+$_e$|A (N/Pr$_e$ ⟨*kanen*⟩)	+IntP$_{af}$ (IntV$_m$)	±A$_{cr}$|$_{fc}$ (Pr$_{fc}$ ⟨*ale*⟩)	...]
(l-j4)	~	[...	+IntP$_{af}$ (IntV$_m$)	+A$_{cr}$|$_{fc}$ (Pr$_{fc}$ ⟨*ale*⟩)	+$_e$|A (N/Pr$_e$ ⟨*kanen*⟩)]

focused activity as distinct from transitive. The focus signal *M* appears here initial in the word.[12]

By comparing (j2) with (j3) we note that the cross-referent pronoun after the verb (that is, $A_{cr}|_{fc}$) may be included or be left out of the intransitive clause. (This makes it somewhat different in functional load from the obligatory A_{cr} in transitive g-f4.) In each instance, however, the arrow indicates that the slot is in focus whether filled or not; when filled, it shows the focus-complement set of pronouns.

I have used the sign # exclusively for the focus-complement third person, which occurs only as zero. The absence of third person plural cross referent is shown by parentheses around the dots in (j2). This is not zero in the same sense, since the optional focus complement *ale* is present there in (j3). Since we do not wish to change direction of focus during the optional absence of *ale*, the arrow points to the empty slots.

Note the ambiguity between (j1″) and (j4″): '(The) cat← cries $_{af}$' versus '(It) ←, (the) CAT, cries $_{af}$'. The third singular in (j4″) $A_{cr}|_{fc}$ – but not the same tagmeme in third plural form in (j4) – is zero, opening the door to this ambiguity.

The emic paradigm for intransitive is seen in (k).

Intransitive clause in tagmemic-notation paradigm (k)

$$\begin{array}{ll|lll} \text{(k-j1)} & \mathrm{IntC}_{af} & \{.. & +\mathrm{IntP}_{af} & +A|_{fc}\} \\ \text{(k-j2–4)} & {}_{ea}\mathrm{IntC}_{af} & \{+_{e}|A & +\mathrm{IntP}_{af} & \pm A_{cr}|_{fc}\} \end{array}$$

The optional nature of $A_{cr}|_{fc}$ is symbolized by ± in the emic representation of (k-j2–4). (Obligatory occurrence of other tagmemes is shown by +.) This freedom is limited to the construction variant which carries the emphatic tagmeme in normal initial position. The restriction will enter the notation only when the etic variants of (k) are shown in (l), p. 209.

In (l) two different kinds of variants require symbols. Whereas the EXPANSION VARIANT of (j2–3), seen with its optional addition, is simply shown by the ± in (l-j2–3), a TRANSFORM VARIANT requires an alternant formula as in (l-j4). Variants of constructions, whether of the expansion or the transform type, can be handled as alternants of which neither has priority over the other.[13] Or one variant of a pair may be treated as a base or norm from which the other is derived by a rule. Derivation of the transform variant (l-j4) from a clause base (l-j2–3) requires the rule that the initial emphatic tagmeme complex be transposed to clause-final position and

[12] Either because of morphophonemic changes due to the particular phoneme sequence beginning the stem, or because of some complexity of relation between intransitive and stative relationships not discussed in Dean's article but perhaps implied by Abrams. (See footnote 7).

[13] Compare similar alternative treatments of allomorphs and allophones.

that the sign ± be changed to + before $A_{cr}|_{fc}$. The base-plus-rule approach, however, makes it difficult to display paradigms and matrices of well-defined units and their variants.

6 Noun phrases

In Dean's article there is little treatment of Bilaan phrases. He has supplied me with a few further forms, which reveal an interesting interlocking between some components of the clause paradigms and certain noun phrases which have one noun as head and another noun (or pronoun) as possessor. In (m) and (n) these are shown in citation and in tagmemic-notation paradigms (Po possessor, NPh noun phrase, H head).

Possessor noun phrase in citation paradigm (m)

(m1)	PoNPh	…	*ungeh*	*-la*	'…	rat	their'
(m1′)		…	*ungeh*	*kuku?*	'…	rat	cat's'
(m2)	$_e$PoNPh	*DALE*	*ungeh*	…	'THEIR	rat	…'

Possessor noun phrase in tagmemic-notation paradigm (n)

(n–m1)	PoNPh	[···	HNPh (N)	Po (N/Pr$_{uf}$A⟨*-an*⟩)]
(n–m2)	$_e$PoNPh	[$_e\vert$Po (N/Pr$_e$ ⟨*kanen*⟩)	HNPh (N)	···]

In (n-m1) the set of regular pronominal possessors of an item is the same set as the non-focus-complement actor of (h-f2). The emphatic pronominal set of (n-m2) is the same as the actor in (h-f4) or the goal in (h-f7). The order relation of pronoun Pr$_e$ and of pronoun Pr$_{uf}$A to the head of the noun phase is the same – before and after respectively – as it is to the predicate of the clause.[14]

An emphatic tagmeme occurs in the possessor noun phrase (n-m2). Our notation $_e|$ is the same as for the substantives in the actor and goal tagmeme complexes of the clause, and shows that our analysis treats this tagmeme as identical with the one that occurs simultaneously with such an actor or goal (in h-f4 and h-f7, for example). The emphatic tagmeme, then, occurs on a par both with a clause-tagmeme string and with a noun-phrase string; it functions on two different structural levels.

Sharply different in character are noun phrases with a marker *di*. These have some relationship to referent function translatable as 'to me', 'for me', and the like. From Abrams' material I gather that the particle *di* occurs when the noun phrase is not functioning as a focus complement. Similarly, Dean's material has a few sentences in which a particle *i* occurs with a noun functioning as actor or goal. Abrams shows this particle as occurring optionally.

---

[14] This fact adds weight to our treatment of focused actor (or goal) as a focus complement to the predicate rather than as a central topic or as a center of focus itself. It suggests, further, that we have here a structural argument from noun-phrase analogy for treating the predicate as the nucleus of the clause with both actor and goal as in some sense marginal or lateral to it.

Many Philippine languages, however, use a particle to mark any noun phrase as a focus complement. In Maranao, *so* marks a phrase as focus-complement subject, focus-complement object, or focus-complement instrument (see McKaughan, 1958, p. 19; 1962). When functioning as a non-focus-complement, a Maranao subject noun phrase takes *o*, an object takes *sa*, a referent takes *ko* (with co-occurrent subjective or objective voice – i.e. actor- or goal-focused activity in the verb) or *sa* ~ *ko* (with co-occurrent instrumental voice), an instrument takes *ko* (with co-occurrent objective or referential voice) or *so* (with co-occurrent instrumental voice). In Tagalog, each of the voices takes as an obligatory focus-complement marker the particle *ang* (see Wolfenden, 1961, p. 16, Table 5). Markers of this type in Maranao and Tagalog prevent much of the Bilaan ambiguity of (i-f3) and (i-f5). In languages where focus markers occur regularly with noun phrases, the order of tagmemes in the clause does not need to be as rigid as it appears to be in Bilaan.

7 Bilaan and English

In view of the widespread activity-focus system, with three or four voices in many Philippine languages, it is not surprising that Bilaan, as Abrams suggests, has a restricted contrast between referent-focused and goal-focused activity. For certain verbs, he says, the *n*- marker in the verb takes over a function of signalling reference-focused activity, with a zero marker for goal-focused activity. This is shown in (o-p) (Abrams, 1960, p. 10). I mark the constructions with an asterisk because the difference in dialect source between Dean and Abrams may make the illustration inapplicable to the Bilaan dialect which Dean describes.

**bat-an ale*.	'He throws them'	(o)
**n-bat-an ale*	'He throws (to) them'	(p)

The importance of this material to us is that it helps to point up one difference between English and many languages of the Philippines. Whereas English has only an active and a passive voice, some Philippine languages have several voices. For this reason we tentatively added the referent clause type to the matrix of (a), even though there was not enough material to add it to our syntax paradigm. Certain of the formal differences between English active and passive clauses and the Bilaan clauses with activity focused toward actor, goal, and referent are symbolized in (q) and (r).

English voices (q)

(q1) $S_a\leftarrow$ $\rightarrow P_{af}$ G
(q2) $S_g\leftarrow$ $\rightarrow P_{gf}$ Agent

Bilaan voices (r)

(r1) P_{af} $\rightarrow A_{fc}$ G
(r2) P_{gf} A $\rightarrow G_{fc}$
(r3) $*P_{rf}$ A G $\rightarrow R_{fc}$
(r4) $\sim *P_{rf}$ A $\rightarrow R_{fc}$ G

Comparing (q1) with (q2) we affirm that English S_a is a subject-as-actor tagmeme differing from homophonous S_g, a subject-as-goal tagmeme. The two are members of a distribution class of homophonous subject tagmemes (see Pike, 1954, p. 131; 1967, p. 245), contrasting in transformations (S_g tagmeme of passive clause transforms to G tagmeme of active; S_a of active clause transforms to agent of passive) and in semantic relations. What holds them together in an emic class of tagmemes is their position before the predicate and their tie to the predicate in person and number (*The boys run: The boy runs*). This tie is symbolized by the arrow pointing right, to the predicate; the arrow pointing left, to the subject, implies the component of observer focus on actor-oriented or goal-oriented activity.

In (r1–4) the Bilaan differences show up clearly. No single distribution class of subject containing actor, goal and referent tagmemes can be fruitfully postulated; different distributional positions are retained for actor and for goal. No concord of person and number links the predicate with the focus complement. The tentative listing (*) of the referent clause type in (r3) implies a more complex voice system for Bilaan than for English. The referent alternant in (r4) allows the focus complement to be drawn back closer to the predicate – provided, however, that this alternant is not used if the two substantives are members of the same subclass, so that their use might result in ambiguity.[15]

---

[15] Thus, Abrams indicates (1960, p. 11) that there is variation in the order of *kayu* 'tree' and *balo* 'bolo knife' in *k-an-lang-gu kayu balo*:*k-an-lang-gu balo kayu*. The degree to which variation occurs in that dialect, however, is not clear from his material. Possibly a special order sometimes occurs when one of two (or more) substantives is a noun and the other a pronoun.

References

ABRAMS, N. D. (1960), *The Verb Complex in Bilaan*, unpublished master's thesis, University of Washington.

DEAN, J. C. (1955), 'The phonemes of Bilaan', *Philippine J. Sci.*, vol. 84, pp. 311–22.

DEAN, J. C. (1958), 'Some principal grammatical relations of Bilaan', *Oceania Linguistics Monogr.* 3, pp. 59–64.

HEALEY, A. (1958), 'Notes on Yogad', *Oceania Linguistics Monogr.* 3, pp. 77–82.

HEALEY, P. A. (1960), *An Agta Grammar*, Manila.

McKAUGHAN, H. (1958), *The Inflection and Syntax of Maranao Verbs*, publication of the Institute of National Language, Manila.

McKAUGHAN, H. (1962), 'Overt relation markers in Maranao', *Language*, vol. 38, pp. 47–51.

PICKETT, V. B. (1960), *The Grammatical Hierarchy of Isthmus Zapotec*, language dissertation no. 56, Baltimore.

PIKE, K. L. (1954, 1955, 1960), *Language in Relation to a Unified Theory of the Structure of Human Behavior*, Parts I, II and III, Summer Institute of Linguistics, 2nd ed. (1967), Mouton.

PIKE, K. L. (1962), 'Dimensions of grammatical constructions', *Language*, vol. 38, pp. 221–44.

WOLFENDEN, E. (1961), *A Re-Statement of Tagalog Grammar*, Manila.

15 P. Healey

Teleéfoól Quotative Clauses[1]

P. Healey, 'Teleéfoól quotative clauses', *Linguistic Circle of Canberra Occasional Papers*, vol. 3, 1964, pp. 25–34.

1 Structure of the quotative clause

One of the contrastive clause types of Teleéfoól is the QUOTATIVE.[2] It is characterized by the obligatory occurrence of some form of the verb *akankalin* as PREDICATE, and the obligatory occurrence of a QUOTE, a clause-level unit unique to the Quotative Clause. The Quote consists of a complete utterance terminated by one of the almost obligatory QUOTATIVE MARKERS, *kalaá* or *oó*. This utterance may be a single-clause or multi-clause utterance, or an exclamation, i.e. a non-clause.[3] The verb *akankalin* and the markers *oó* and *kalaá* do not occur elsewhere.

The clause-level units occurring in the Quotative Clause are limited to three: optional Subject, obligatory Quote, and obligatory Predicate (a form of *akankalin*), in that order. The Naming sub-type alone has an additional obligatory unit, an Object, which follows the Subject in order.[4] Most other clause types contain many more (optional) clause-level units, including Time, Accompaniment, Location, Beneficiary, Complement, Destination, Instrument.

The Quotative Clause is in fact a class of clause types which resemble one another much more than they resemble any other clause type. However, they also differ contrastively from one another in several important ways, and so may be said to constitute sub-types of the Quotative Clause. The following sub-types have been distinguished.

---

[1] This is the language spoken in the vicinity of Telefomin, Sepik District, New Guinea.

[2] This paper represents findings based upon six hundred examples of the Quotative Clause found in one hundred and twenty pages of transcribed text (including a group discussion, a conversation, narratives and discourses on various subjects, sermons and prayers, and involving at least ten different speakers belonging to two generations, two speakers predominating), sixty pages of translated material, and elicited data.

[3] A clause occurring as Quote is characterized by the Quotative Marker *oó* or *kalaá* that follows it, in contrast to an included clause within any other clause-level unit, which is obligatorily marked by a following pronoun.

[4] In clause types other than the Quotative, the Object is an optional unit.

(a) SAYING, in which the final Predicate of the Quote may contain any tense, and the Quote Marker is *oó*, or, more rarely (16 per cent of occurrences), one of the common Utterance Terminals, such as the Indicative Indicators *koo*, *kwa*, or *kuba*, or the Interrogative Indicators *a* or *áka*. These occur more frequently in this context in the speech of members of the older generation. The younger generation prefer always to use *oó* rather than to distinguish the character of the Quote by using Utterance Terminals. Only the Saying sub-type may have an exclamation as Quote.

(b) DESIDERATIVE, in which the final Predicate of the Quote contains a non-real tense (Potential or Future), and the Marker is obligatorily *oó*.

(c) DIRECT IMPERATIVE, in which the tense of the Quote-final Predicate is usually Potential, but may be Dependent or Abilitative, the Marker is *oó*, and the Quotative Clause Predicate is unmarked for person and followed by an Indicative Utterance Terminal (*akeeta koo*). That is, it is a non-finite closed Predicate. Only the Imperative sub-type may not have an Equational Clause as Quote. It should be noted that this is the Direct Imperative ('Do it!'), not the Quoted Imperative (' "Do it!" he said.'), which is a variety of the Saying sub-type.

(d) NAMING, in which the Quote consists of the NAME (analysable as an Equational Clause), and an obligatory Object occurs. This Object is a Noun Phrase, and is the thing named or REFERENT. A clause-level free Object does not occur in other sub-types of Quotative Clause. The Marker is *oó*.

(e) PERCEPTIVE, in which the Potential tense does not occur in the Quote-final Predicate, and the Marker is obligatorily *kalaá*.

The contrastive features of these five Quotative Clause sub-types are tabulated in Table 1, and illustrated in a syntax paradigm in Table 2.

It will be noted that the Potential tense may occur in the Quote of the Saying sub-type. In the majority of these cases the Potential represents a Quoted Imperative, and there is little ambiguity with the Desiderative sub-type. However, a few utterances occur involving other than second person, particularly first person cases, and here the Quotative Clause is ambiguously of the Saying or the Desiderative sub-type.[5]

---

[5] When there is an Associated Clause (see section 3 below), three factors operate to resolve this ambiguity: (a) the verb occurring as Predicate of the Associated Clause (*bákaamin* and a very few others occur in the Saying sub-type, while any verb may occur in the Desiderative); (b) the Subject of the Associated verb is the same as that of *ankankalin* in the Saying sub-type, but may be different in the Desiderative; (c) the connective suffix is predominantly *-ileé* for the Saying sub-type, predominantly *-ta* for the Desiderative.

Table 1 **Contrastive features of quotative clauses**

Clause sub-type	*Object*	*Quote* *Tense of verbal predicate*	*Marker*	*Predicate and clause terminal*
Saying	–	non-potential, potential	*oó*, ⟨*koo*⟩	finite; non-finite open
Desiderative	–	potential	*oó*	finite; non-finite open
Direct imperative	–	potential	*oó*	non-finite closed
Naming	noun phrase	(equation) name	*oó*	finite; non-finite open
Perceptive	–	non-potential	*kalaá*	finite; non-finite open

Table 2 **Illustrative paradigm of quotative clauses**

Clause sub-type	*Object*	*Quote* *Predicate*	*Marker*	*Predicate and clause terminal*	*Free translation*
Saying	–	*únbí* 'I went'	*yoó*	*akeélá koo* he said	He said he went
Desiderative	–	*unoón* 'I'll go'	*oó*	*akeélá koo* he wanted	He wanted to go
Direct imperative	–	*unoón* 'I'll go'	*oó*	*akeeta koo* must	I must go
Naming	*boómi wín* her name	*Fuúmeen* 'Fuumeen'	*oó*	*akeéla koo* he called	He called her Fuumeen
Perceptive	–	*únbú* 'She went'	*kalaá*	*akeélá koo* he saw	He saw her go

2 Meaning of the quotative clause

The semantic range of the verb *akankalin* extends from 'say, think' (Saying sub-type) to 'see, know, feel' (Perceptive sub-type).[6] However the whole class of clauses has been termed Quotative because all must be regarded as a kind of 'direct speech' – or 'direct cerebration'. The pronoun suffix of the final verb of the Quote is that which one would predict for direct speech involving a SPEAKER (Subject of *akankalin*), an ACTOR (Subject of the final Predicate of the Quote), and a HEARER.

The Hearer is often not overtly indicated. It may, however, be indicated by a benefactive affix in the Predicate of an Associated Clause: or, very occasionally, when there is no other indication in the context at all, by a first or second person prefix on *akankalin* itself. For example, the Saying example in the paradigm could read: *únbí yoó nakeélá koo*. 'He told me he went.' (' "I went," he said to me.').

If Actor and Speaker refer to the same individual, or if the Actor includes the Speaker, the Actor will be a first person suffixed pronoun. 'He told me he'd go' may be literally translated, '"I'll go," he told me'. 'He wanted to go' may similarly be translated, '"I will go," he wanted'.

If Actor and Hearer refer to the same individual, the Actor will be a second person suffixed pronoun. So, 'He told me to go' would be literally, '"Go!" he told me'. 'He wanted you to go' would be, '"He'll go", he wanted (concerning you)'.

If Actor, Speaker and Hearer all refer to different individuals, then the Actor will be a third person suffixed pronoun. So, 'He told them I'd gone' would be. '"He has gone," he said to them (concerning me).' 'He wanted me to go' would be, '"He will go," he wanted (concerning me)'. 'He told them you'd gone' would be, '"He has gone," he told them (concerning you)'. 'He saw them go' would be, '"They went," he saw'.

The Quotative Clause sub-types appear to have roughly the following central areas of meaning:

Saying: 'say, tell (him),[7] think, wonder' (The Quote is what is said, etc.)
Desiderative: 'want to, try to, be going to' (The Quote is what one wants to do or wants done.)

---

[6] A single grammatical feature encompasses a similar wide semantic range in certain other known languages. In the Turkic languages, especially the Central Asian languages such as Uzbek, one construction covers saying, wishing, assuming, purposive, naming and perceptive (information from S. A. Wurm). Similarly, in Waiwai of British Guiana one construction covers saying, naming, perceptive and many related concepts (Hawkins, 1962).

[7] The brackets ( ) around a particular word denote the use of that word to represent the whole syntactic class to which it belongs.

Direct imperative: 'do!, must, should, ought to, let (him)' (The Quote is the command, etc. All person suffixes occur in the Direct Imperative.)
Naming: 'call (him), name (him)' (The Quote is the name.)
Perceptive: 'see (that), know (that), feel that, realize that' (The Quote is what is seen, etc.)

The Saying sub-type particularly has a very wide semantic range of usage, as illustrated by the following, of which the first three have an Exclamation as Quote, the next two an Equational Clause as Quote, and the final two a Predicative Clause as Quote.

(*uu*) *oó akeélá koo.* 'He agreed.' ('"Yes," he said.')
waák(*uu*) *oó akeélá koo.* 'He refused.' ('"No," he said.')
mísám oó akeélá koo. 'He thanked (him).' ('"Thanks," he said.')
aáfeén oó akeélá koo. 'He believed it.' ('"It is true," he said.')
boótá támbálim oó akeélá koo. 'He approved it.' ('"That is good," he said.')
unoónteemaálú tab oó akeélá koo. 'He doubted that she would go.' ('"She will not go perhaps," he said.')
Fuúmeen kanúbú yoó akeélá koo. 'He accused Fuumeen.' ('"Fuumeen did it," he said.')

The areas of meaning of the various clause sub-type overlap to some extent, and make analysis more difficult. Within each area of meaning, the exact meaning of the clause is specified by the context, especially by the Associated Clause, if present, and the content of the Quote, as can be seen in the above examples. Thus, the Saying sub-type means 'ask' if the verb *dákaákaamin* is the Predicate of the clause associated with the Quotative Clause, and/or if the Quote is an interrogative utterance.

3 Associated clause

In about one-third of its occurrences the Quotative Clause has associated with it another clause in apposition with it. In most instances the Associated Clause precedes the Quotative Clause, in some instances it follows, and in about as many instances Associated Clauses both precede and follow a Quotative Clause, redundantly conveying similar information.

The main function of the Associated Clause is to spell out the specific meaning of the *akankalin* verb of the Quotative in any particular instance. For instance, the commonest verb of the Associated Clause of a Saying Clause is *bákaamin* 'say', while the commonest one with a Perceptive Clause is *utámaamin* 'see'. The Associated Clause also serves as a vehicle for additional clause-level units that do not occur in the Quotative Clause, such as Time and Beneficiary.[8]

In the great majority of instances there is a connective suffix on the Predicate of the first of the two apposed clauses, and this is overwhelmingly either *-ta* or *-ileé* (~ *-leé* after a vowel), of which *-ta* is the commoner. No meaning distinction between these connectives can be ascertained, but they differ slightly in distribution, as will be demonstrated later.

The verb occurring as Predicate of the Associated Clause, therefore, and to some extent its connective suffix, provide extra diagnostic criteria for establishing the sub-types of Quotative Clause, which have already been set up on internal structural evidence. There is a further criterion for distinguishing the Saying Clause from the Desiderative Clause. The suffixed Subject of the Predicate of the Associated Clause is always the same as that of the Quotative Predicate *akankalin*, except in the case of the Purposive variant of the Desiderative sub-type, where it may be the same or different.

The SAYING Clause co-occurs most frequently with the verb *bákaamin* 'say, tell (him)', often co-occurs with *dákaákaamin* 'ask' and *fúkúnin* 'think', and less frequently with certain other verbs. The connective is usually *-ileé* (55 per cent of all occurrences), though *-ta* also occurs. The Saying Clause in the paradigm could therefore be extended to: *bókoonalaleé únbí yoó akeélá koo*. 'He said he went.' ('He said "I went" he said.')

The DESIDERATIVE may have an Associated Clause. When this occurs the pair of clauses have a PURPOSIVE meaning, the Quote being the purpose for which the action of the verb of the Associated Clause is performed. The Associated Clause may contain any verb as Predicate. This Associated Clause is not in an apposition relationship with the Quotative Clauses as are the Associated Clauses of the other sub-types. A special case of the Desiderative (Purposive) is that involving a 'why' clause as Quote. The 'why' clause has only two possible forms, *íntabeén* 'Why (is it)?' (Equational), and, *íntabeén núlán* 'Why did I/you/he?' or, *íntabeén núlúm* 'Why did we/you/they?' (Predicative). The suffixed pronouns so far observed in this situation are always first person. The Associated Clause always follows a 'why' Quote Clause. The most frequent connective occurring with the Desiderative (Purposive) is *-ta* (70 per cent of all occurrences), but *-ileé* also occurs.

---

[8] The first and second person singular Beneficiaries may occur at the morphological level within the Quotative Clause as a prefix, *na-* and *ka-* respectively, to *akankalin*, forms which in all other prefixable verbs indicate Object. However, where there is no ambiguity the third person and plural form *a-* is used even when the Beneficiary of the Associated Clause is first or second person singular. *na-* and *ka-* therefore occur relatively rarely with *akankalin*.

íntabeén núlán oó akeeta tálbá ya. 'Why did he come?' ('"What will I do?" wanting he came.')

talnalata unoón oó akeélá koo. 'He came because he wanted to go.' ('He came "I will go" he wanted.') (This is an extension of the Desiderative clause in the paradigm.)

The IMPERATIVE type does not have an Associated Clause.

The NAMING type may have an Associated Clause, in which *bákaamin* 'say' is the only verb observed to occur. *-ileé* and *-ta* occur, but there are insufficient examples to assess their relative frequency. An extension of the Naming Clause in the paradigm could be:

bókoonalaleé boómí wín Fuúmeen oó akeélá koo. 'He called her Fuumeen.' ('He said her name "Fuumeen" he called.')

The PERCEPTIVE type has an Associated Clause in at least half of its occurrences, and in 70 per cent of these occurrences the verb occurring in the Associated Clause is *utámaamin* 'see, know, feel'. *feenin* 'wait, watch for' and *tinángkaamin* 'hear' also occur. There is a connective suffix *-tiyeé* which occurs with the Perceptive in 40 per cent of its occurrences, and rarely occurs elsewhere. The connective *-ta* also often occurs here (30 per cent of occurrences); *-ileé* occurs more rarely. The Perceptive Clause in the paradigm might be extended thus:

utámatiyeé únbú kalaá akeélá koo. 'He saw her go.' ('He saw "She went" he saw.')

4 Complex quotative constructions

Two types of complex Quotative construction occur in the language – SEQUENTIAL and NESTED constructions.

The SEQUENTIAL construction involves a string of Quotes, each ending in *oó* (or *kalaá*), occurring before a single *akankalin* Predicate. In this construction, the Quotes are in a coordinate relationship with one another, and therefore all belong to the same sub-type. For example, *talbaálú kalaá únbú kalaá akeélá koo.* 'He saw that she didn't come, she went.' ('"She didn't come, she went," he saw.')

The NESTED construction on the other hand may consist of various different Quotative Clause sub-types. A number of Quotative Clauses may occur in a layered or nested arrangement such that the first span of speech containing a Quote and terminating in a form of the verb *akankalin* followed by *oó* or *kalaá* is the Quote of the second *akankalin*; that Quote plus the second *akankalin* followed by *oó* or *kalaá* is the Quote of the third *akankalin* and so on. The following example contains four syntactic layers, all of clause level, nested within one another. The innermost is a Quote consisting of a minimal clause, the Predicate *kanumanteémí* 'I will do it' plus *kalaá*; the second layer, *kanumanteémí kalaá akeébá* '"I will do it," he has decided', a Perceptive Clause, plus *kalaá*, is the Quote of the third layer, also a Perceptive Clause. Similarly, the whole third layer plus *oó* is the Quote of the fourth and outermost layer, a Saying Clause.

Transitive	*kanumanteémí*		
clause	I will do it		
Perceptive	*kanumanteémí kalaá*	*akeébá*	
clause	'I will do it,'	he has decided	
Perceptive	*kanumanteémí kalaá*	*akeébá kalaá*	*akeebaálib*
clause	'He has decided to do it,'		you don't know
Saying	*kanumanteémí kalaá akeébá kalaá akeebaálib oó akeenalaleé*		
clause	'You don't know that he has decided to do it,' he said ...		

Reference

HAWKINS, R. E. (1962), 'Waiwai translation', *Bible Translator*, vol. 13, pp. 164–71.

16 D. G. Hays

Dependency Theory: A Formalism and Some Observations

D. G. Hays, 'Dependency theory: a formalism and some observations', *Language*, vol. 40, 1964, pp. 511–25.

1 Introduction

In the general theory of linguistics, DEPENDENCY THEORY competes for the same role as IMMEDIATE-CONSTITUENT (IC) THEORY: each formalizes a view of tactic relations among elementary units in utterances. In each theory, a finite apparatus is taken to CHARACTERIZE (or GENERATE) a set of utterances that may be infinite. In addition, each theory specifies a structure, or perhaps alternative structures, for each utterance it generates.

The essential insight formalized by IC theory was stated clearly by Harris in the article 'From morpheme to utterance' (1946). There are composite expressions (phrases) that have the same distribution as minimal syntactic units; for example, there are noun phrases that have the same distribution as some nouns. This insight leads to a principle of empirical research, since the integrity of a phrase in an utterance is confirmed by the existence of a minimal unit substitutable for it. This principle is not sufficient, for two reasons. In the first place, some phrases have distributions that differ from all those of minimal units. Secondly, there are infinitely many varieties of mutually substitutable stretches of utterances, and therefore the principle does not lead to discovery of a finite characterization. Nevertheless, the principle is significant for practical linguistics. The same insight leads also to a formal calculus of phrase types or classes, namely Chomsky's theory of context-free phrase-structure grammars (1956). The structures assigned by grammars of this type will be called P-markers. The importance of IC theory and the use of P-markers in linguistics are entirely independent of the empirical principle of substitution; characterization theories and research methods (so-called discovery procedures) must be studied separately.

The insight formalized by dependency theory is that particular occurrences of minimal units are directly related to one another. Occurrence of an adjective before a noun, for example, is more closely related to the noun occurrence than to anything else. In describing the distribution of a class of adjectives, therefore, one naturally refers to the classes of nouns with which the adjectives occur. Similarly, an occurrence of a preposition is generally most directly related to two other occurrences, its object and the unit modified by the prepositional phrase. Prepositions are therefore regarded as connectors, and their distributions are most easily stated by reference to classes of objects and classes of governing or modified units. This kind of observation about language is sometimes expressed in terms of the VALENCES of units for one another (Iordanskaja, 1963).

Some of Hjelmslev's empirical principles are closely related to the insight behind dependency theory (Hjelmslev, 1961, p. 24), but empirical dependency in his sense cannot be identified with abstract dependency in the sense of the present paper, since he explicitly differentiates dependencies from other kinds of relations, whereas the present theory intends to be complete, i.e. to account for all relations among units of utterances.

Dependency theory is actually a characterization theory, not necessarily associable with any empirical method or principle. It is a theory of grammars, with abstract mechanisms for characterizing sets of utterances and for assigning to them certain structural descriptions, which will be called D-trees.

The present paper is an introductory account of the theory, and a survey of its current state of development. Section 2 presents a formalism for the theory, identifying the components of any dependency grammar. Next follows the description of a generative procedure (section 3) and some decision procedures (section 4). These are alternative and equivalent explications of the relation between finite grammars and infinite sets of utterances. The structures assigned by dependency theory are described in section 5. IC theory and dependency theory are compared in section 6, which also gives some attention to their ability to assign similar structures to utterances in a given set. Section 7 discusses extensions of dependency theory by addition of transformations or by embedding in a multistratal system. Finally, section 8 examines the problem of empirical application.

2 Formal dependency theory

In the more familiar IC grammars, two alphabets are used, terminal and non-terminal (Chomsky and Miller, 1963). One interpretation makes the terminal alphabet a lexicon or morpheme list. The non-terminal alphabet is a set of names or symbols for phrase types, perhaps including phrases of a single element.

A dependency grammar also uses two alphabets, and one is TERMINAL in exactly the same sense: the utterances to be characterized are strings of elements of the terminal alphabet. The non-terminal AUXILIARY alphabet is a set of names or symbols for types of occurrences of terminal-alphabet characters. Each element of the terminal alphabet is a minimal syntactic unit – for example, a morph, morpheme, or word. The theory allows minimal syntactic units to be syntactically ambiguous: a given morpheme can act as a verb stem in one utterance, as a noun stem in another. An auxiliary character is, in this sense, unambiguous; a terminal character can correspond to any number of different auxiliaries, each with fixed syntactic properties. On the other hand, an auxiliary character can correspond to a long list of different terminals, namely, all those that can occur with its syntactic properties. Bar-Hillel has called the many-many correspondence between auxiliary and terminal elements the ASSIGNMENT FUNCTION of a grammatical system (Bar-Hillel, Perles and Shamir, 1961).

A DEPENDENCY RULE is a statement about the valence of one kind of syntactic unit. It consists of one auxiliary character as GOVERNING element, and any finite number of auxiliary characters as DEPENDENT elements; terminal characters never appear in dependency rules. A convenient notation (due to Gaifman, 1965) is

$$X_i(X_{j_1}, X_{j_2}, \ldots, *, \ldots, X_{j_n}). \tag{1}$$

Here X_i is the governing element; $X_{j_1}, \ldots, X_{j_n}$ are dependent elements, n in number. The rule implicitly orders all of the elements in it, by the order of the elements within the parentheses. The asterisk marks the position of the governor among its dependents. The valence of X_i given by this rule is just the set of dependent elements stated. The order given to the elements by the rule is the relative order of corresponding terminal characters in utterances, but other terminals can be interspersed among them.

A dependency rule can be regarded as a statement that an occurrence of a syntactic unit corresponding to X_i can govern simultaneously occurrences of units corresponding to X_{j_1}, to X_{j_2}, and so on up to X_{j_n}. As an example, take the hypothetical English rule

$$V_\alpha(N_{pl}, *, N, D_\beta),$$

where V_α is a class of verb morphemes, N_{pl} a class of plural nouns, N a noun class, and D_β a class of adverbs – say, of manner. This rule could be used in connection with utterances like *children eat candy neatly*.

The last example illustrates the nearest approach in dependency theory to classification of PHRASES. In general, minimal syntactic units are the elements of dependency calculations, but whenever necessary an auxiliary character like N_{pl} can be introduced. Interpreting the minimal units of the theory as morphemes, we must understand N_{pl} as corresponding to occurrences of noun-stem morphemes governing plural morphemes. In this fashion, a dependency grammar can take into account distant influences such as agreements between subject and predicate complement of a copulative verb.

STATE INSTITUTE OF EDUCATION * of West Bengal

Formally, a dependency grammar must contain two other kinds of rules. A rule of the form

$$X_i(*) \tag{2}$$

states that a terminal corresponding to X_i can occur without dependents. Likewise, a rule

$$*(X_i) \tag{3}$$

states that a unit corresponding to X_i can occur without being governed. Such a unit is the CENTRAL or MAIN element of the utterances in which it so occurs.

Thus a dependency grammar consists of a terminal alphabet, an auxiliary alphabet, an assignment function linking the two, and a set of rules including some of each of the three forms specified. The utterances characterized by the grammar are strings over the terminal alphabet. The assignment function matches terminal characters with syntactic types. A rule of form (3) marks the center of an utterance, rules of form (2) mark the outer limits, and rules of form (1) mark the valence or dependency relations among the units of the utterance. How such a grammar characterizes a set of utterances is described in sections 3 and 4.

3 Generative characterization

An utterance is a finite string of minimal syntactic units. A grammar characterizes a set of utterances; as Chomsky rightly argues, a collection of statements about a language that does not characterize its utterances does not deserve to be called a grammar (Chomsky and Miller, 1963). (Incidentally, this argument has the terminological consequence that no special descriptive term is needed for the class of grammars that do characterize sets of utterances, there being no grammars that do not.)

Classically, mathematical logic offers at least two approaches to the characterization of sets: generative functions and decision procedures. Postal is mistaken in arguing that the former approach is logically more fundamental than the latter (Hamp, 1963, p. 36), and Chomsky is almost equally wrong in asserting that the former is overwhelmingly predominant in the history of linguistics (Chomsky and Miller, 1963). Chomsky is correct, however, in differentiating between (a) the generation-decision distinction in characterization theory and (b) the production-recognition distinction in theories of behavior or, as Chomsky puts it, performance (1963). Although the distinctions are different, the difference is susceptible of overstatement; a practical parsing device for a class of sets of utterances is *a fortiori* a decision procedure for the class.

A class of grammars was defined in section 2, but only up to the point of a mechanism for generating utterances or a decision procedure. The first kind of abstract mechanism, given a dependency grammar, generates an infinite list of strings: those of the language characterized by the grammar. A decision procedure, given the grammar, decides for each string over the terminal alphabet whether it does or does not belong to a certain set; that set is also said to be the language characterized by the grammar. Naturally, it is essential to prove, for a CLASS of grammars, that the set generated by a certain procedure is the same as the set accepted by a given decision procedure whenever the same grammar is supplied to both.

The present section describes a generative procedure; section 4 presents a decision procedure.

The generative procedure has two stages; the first stage is a sequence of steps, with all except the first carried out on the same plan. The first step is the choice of a central element. For this step, a rule of form (3) is chosen; it contains a single auxiliary element, X_i. The second step is to choose a rule of form (1) or (2) with the same X_i as governing element. If the rule chosen has the form $X_i(*)$, the first stage of the procedure is complete. Otherwise, the rule has dependent elements $X_{j_1}, X_{j_2}, \ldots, X_{j_n}$. Take the first and choose a rule with X_{j_1} as governing member. The new rule may specify dependent members; if it does, they are added to the object being constructed. The construction continues until a rule has been chosen for each character introduced.

Certain information must be recorded each time a rule is chosen. A notation is the following. Let $*(X_i)$ be the first rule chosen, and $X_i(X_{j_1}, X_{j_2}, \ldots, *, \ldots, X_{j_n})$ the next. Let $X_{j_1}(*)$ be the rule chosen for X_{j_1}. Then, after the first choice, record

$*(X_i)$.

After the second,

$*(X_i(X_{j_1}, X_{j_2}, \ldots, *, \ldots, X_{j_n}))$.

After the third rule choice, record

$*(X_i(X_j(*), X_{j_2}, \ldots, *, \ldots, X_{j_n}))$.

This expression shows clearly that rules have been chosen for X_i and X_{j_1}, but not for X_{j_2} and the rest. When this stage of the procedure is complete, the expression contains a parenthesized term after each auxiliary character, but of course some parenthesized terms are of the form (*).

The second stage of the generative procedure replaces each auxiliary character with a terminal one and orders the terminal characters in a string. The replacement is controlled by the assignment function of the grammar; each auxiliary corresponds to a list of terminals, of which one is chosen.

The order of the units in each dependency rule is given as part of the rule; these various orderings are combined in accordance with a principle shown by the following example.[1]

Let the expression developed with dependency rules be

$$*(X_i(X_{j_1}(X_{k_1}(*), *), *, X_{j_2}(*), X_{j_3}(X_{p_1}(*), *))).$$

The six auxiliary characters must be ordered. The main element, X_i, is governing member in the rule $X_i(X_{j_1}, *, X_{j_2}, X_{j_3})$. Assign order numbers to these four characters in the obvious way, that is, number 1 to X_{j_1}, number 2 to X_i, number 3 to X_{j_2}, and number 4 to X_{j_3}. Now, X_{j_1} governs in $X_{j_1}(X_{k_1}, *)$. Taking the number 1 already assigned to X_{j_1}, expand it: assign 1.1 to X_{k_1} and 1.2 to X_{j_1}. The number finally assigned to X_{j_3} is 4.2, and to X_{p_1}, 4.1. Taking these numbers in order gives

1.1 X_{k_1}
1.2 X_{j_1}
2 X_i
3 X_{j_2}
4.1 X_{p_1}
4.2 X_{j_3}

Thus the sentence is ordered by composition of rule orders.

Any dependency grammar, then, characterizes the set of sentences that can be generated by this procedure, using the rules and assignment function of the grammar. Since all the lists in a dependency grammar are finite, this generative procedure obviously requires only a finite number of finite operations for any finite string that it can generate.

4 A decision procedure

Consider the class of all finite strings of elements of a given terminal alphabet, and suppose that a dependency grammar is specified. A decision procedure must assign each string to one of two sets: one is the set of strings characterized by the grammar, and the elements of this set are said to be ACCEPTED by the decision procedure; the other set consists of strings of elements that are REJECTED by the procedure.

The first step in the procedure to be described here is to list all auxiliary characters corresponding to each minimal syntactic unit of a given string. Next, consider all rules of the grammar that match substrings of auxiliaries. For example, suppose the string under examination consists of four minimal syntactic units, corresponding to auxiliary characters as given in

---

[1] David G. Hays (1961). The procedure was suggested by Thomas Mullikin.

Table 1 **Assignment of auxiliaries to syntactic units**

Minimal syntactic units	1	2	3	4
Corresponding auxiliary characters	X_{11} X_{12}	X_2	X_{31} X_{32} X_{33}	X_{41} X_{42}

Table 2 **Application of one-character rules**

Minimal syntactic units	1	2	3	4
Corresponding auxiliary characters	X_{11} X_{12}	X_2	X_{31} X_{32} X_{33}	X_{41} X_{42}
One-character rules	$X_{11}(*)$		$X_{32}(*)$	$X_{41}(*)$

Table 3 **Application of rules spanning two or three units**

Minimal syntactic unit	1	2	3	4
Corresponding auxiliary character	X_{11} X_{12}	X_2	X_{31} X_{32} X_{33}	X_{41} X_{42}
One-character rules	$X_{11}(*)$		$X_{32}(*)$	$X_{41}(*)$
Two-character rules		$X_2(*, X_{32}(*))$		
Three-character rules	$X_{11}(*, X_2(*, X_{32}(*)))$	$X_2(*, X_{32}(*), X_{41}(*))$		

Table 1. If there are rules $X_{11}(*)$, $X_{32}(*)$, and $X_{41}(*)$ in the grammar, add them to the list (Table 2). Now the following two-character rules, if they exist, match substrings: $X_2(X_{11}, *)$; $X_2(*, X_{32})$; $X_{31}(*, X_{41})$; $X_{32}(*, X_{41})$; $X_{33}(*, X_{41})$; $X_{41}(X_{32}, *)$; and $4_{42}(X_{32}, *)$. Note that a character can be matched as dependent member in a rule, that the order in a rule must match order in the string, and that no jumps are allowed; thus $X_{11}(*, X_{32})$ does not match, since it jumps over unit 2.

Suppose, for the sake of the example, that the only two-character rule in the grammar, among those that could match, is $X_2(*, X_{32})$. Turning to three-character substrings, we note that there are just two spans of three characters, namely, the first, second, and third units and the second, third, and fourth. The first of these matches any of the following rules: $X_2(X_{11}, *, X_{32})$; $X_{11}(*, X_2)$; $X_{12}(*, X_2)$. The latter rules are matchable if we take X_2 as governing X_{32}. The second span matches any of the following rules:

$X_2(*, X_{32}, X_{41})$; $X_{41}(X_2, *)$; $X_{42}(X_2, *)$. No jump is involved in the latter rules, since the third syntactic unit is taken with the second.

Suppose that the grammar actually includes the rules $X_{11}(*, X_2)$ and $X_2(*, X_{32}, X_{41})$. Then the tabulation to date has the appearance of Table 3. A rule covering the whole string of four minimal syntactic units will account for the string in terms of the grammar. The following rules would match: $X_{11}(*, X_2)$; $X_{12}(*, X_2)$; $X_{11}(*, X_2, X_{41})$; $X_{12}(*, X_2, X_{41})$; $X_2(X_{11}, *, X_{32}, X_{41})$; $X_{41}(X_{11}, *)$; $X_{42}(X_{11}, *)$; $X_{41}(X_{11}, X_2, *)$; and $X_{42}(X_{11}, X_2, *)$. The first of these rules would lead to the expression

$$X_{11}(*, X_2(*, X_{32}(*), X_{41}(*))).$$

The third leads to

$$X_{11}(*, X_2(*, X_{32}(*)), X_{41}(*)).$$

Similarly, each rule leads to an expression corresponding to the whole string and built out of subexpressions already tabulated.

If any rule covering the whole string exists in the grammar, the final step is to test whether the first character of the expression obtained appears in a rule $*(X_i)$. If it does, the decision procedure accepts the string; it belongs to the set characterized by the grammar.

The set of strings accepted by this procedure, given a certain grammar, is almost obviously identical with the set generated by the procedure of section 3, given the same grammar. Gaifman proves equivalence for a similar pair of procedures (1965).

The procedure just described is almost (but not quite) an outline of a practical algorithm for recognition, as required, for example, in such branches of applied linguistics as machine translation. It is closely related to procedures suggested by Cocke (Hays, 1962) and Lecerf (1960). The differences are simplifications for the sake of clarity.

5 Structures

The structure of an utterance is a representation of its parts and their properties and interrelationships. It is clear that any structure must reveal the application of a grammar to a particular string, and also that ambiguous utterances must have alternative structures, at least in so far as their ambiguities are grammatical. Other attributes and uses of structures are subject to debate; further remarks appear in later sections. Here it is shown that dependency theory permits assignment of structures to utterances – structures that differ considerably from constituency diagrams.

Dependency structures are in fact represented by the expressions already used in sections 3 and 4. In the example of the previous section, a possible expression for the complete string was the following:

$$*(X_{42}(X_{11}(*), X_2(*, X_{32}(*))), *).$$

According to this expression, the fourth minimal syntactic unit of the string, interpreted as type X_{42} rather than X_{41}, is the syntactically central unit. It governs the first and second units of the string, those units being given syntactic interpretations X_{11} and X_2, respectively. Finally, the second unit governs the third, interpreted as X_{32}.

In an expression of this form, the assignment function is not represented. Replacing each asterisk (except the first) with a terminal character yields an expression which corresponds to exactly one terminal string and shows precisely how the dependency grammar accounts for that string. In cases of ambiguity, two or more such expressions correspond to the same terminal string. For a terminal string that does not belong to the set characterized by the given grammar, no such expression can be written.

Returning to an earlier example, consider *Children eat candy neatly*. An expression for this utterance might be as follows:

$$*(V_\alpha(N_{pl}(\textit{childr}, S(\textit{-en})), \textit{eat}, N(\textit{candy}), D_\beta(\textit{neat}, L(\textit{-ly})))).$$

The terminal string, in this interpretation, is a string of morphs: *childr -en eat candy neat -ly*. How such a string might be converted into a string of phonemes or graphemes is discussed in section 7. For an irregular noun like *children*, assignment of *childr* to a syntactic category N_{pl} is not especially unnatural, but to assign *book* to N_{sg} when it occurs without a plural affix, and to N_{pl} when the affix is present, does seem unnatural. Such double assignments are by no means inherent in dependency theory; in section 7, various techniques for avoiding them are proposed.

The expressions used here for dependency structures are linearized expressions for labeled, rooted trees (Gorn, 1961). A tree is a set of elements on which a relation is defined; in effect, the relation connects every element of the set directly to at least one other, and there is exactly one sequence of connections from any element to any other. A tree is rooted if one element, the origin, is distinguished from the rest. A tree is labeled if properties of its elements – other than the defining relation – are noted (Berge, 1959). In dependency structures, the origin of the tree is the syntactically central unit; the relation is that of direct dependency, and the labels are auxiliary characters and order indicators.

Trees are often presented in diagrams like that of Figure 1; the structure illustrated there is the one represented by the expression just analysed (in the figure, MSU = minimal syntactic unit). The root element of the tree is at the top in the diagram; solid lines stand for dependency connections, the lower element depending on the higher. Each element is labeled explicitly with an auxiliary character, implicitly with an order indicator.

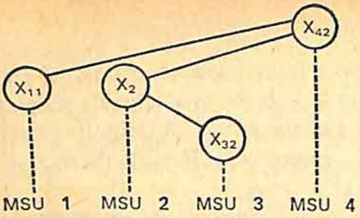


Figure 1 Diagram of a dependency structure

Below the tree is a string of minimal syntactic units, each joined to the corresponding element of the tree; the broken lines map the assignment function. When such a diagram is accurately constructed, there are no intersections of connection lines with one another or with projection lines; this follows directly from the ordering rules given earlier as part of the construction procedure. This attribute, called PROJECTIVITY (Ihm and Lecerf, 1963), is analogous to non-discontinuity of constituents in immediate-constituent theory.

Tree diagrams and linear expressions correspond biuniquely; in passing from one to the other, there is no opportunity for variation. Thus the dependency structure of a sentence can be understood by reference to either without ambiguity.

6 Comparison of dependency and IC theories

Two theories of grammar can be compared on divers grounds. Formally, two issues arise. 1. Can every set of strings characterizable by a grammar of one type be characterized by a grammar of the other type? Two grammars are said to be WEAKLY EQUIVALENT if they characterize the same set of strings; let us say that two theories are WEAKLY EQUIPOTENT if every grammar of either type is weakly equivalent to some grammar of the other type. 2. Can the structures assigned to strings by grammars of the two types be set into correspondence in such a way that corresponding structures are attributed to the same strings? In other words, choose a grammar of one type and consider whether there exists a grammar of the other type that not only characterizes the same set of strings but also attributes to each string structures corresponding to those attributed to the same string by the first grammar. Two grammars are said to be STRONGLY EQUIVALENT if they have this property; let us say that two theories are STRONGLY EQUIPOTENT if every grammar of either type is strongly equivalent to some grammar of the other type.

In addition to these formal questions, there are others. The relative economy of grammars of the two types might be discussed: does one theory generally require shorter lists of rules, or smaller auxiliary alphabets? Intuitive appeal, especially with respect to the form of the structures assigned, is likely to influence the personal choices of investigators. If the theories in question are partial theories of linguistic behavior, to be included in systems with, for example, transformations or multiple strata, one theory may be considerably more appropriate than the other for integration into the larger system. If empirical evidence beyond sentencehood is admitted, still further bases of comparison may be found.

Dependency theory is weakly equipotent to IC theory. The proof is due to Gaifman, and is too lengthy to present here (1965). The consequence of Gaifman's theorem is that the class of sets of utterances characterizable as in sections 3 or 4 is Chomsky's class of context-free languages (1956).

A demonstration of strong equipotence rests on a preestablished formal definition of correspondence between structures. Gaifman studies two definitions. According to his own definition, dependency theory is strongly equipotent to an essentially trivial subclass of IC grammars. The other definition is taken from Hays (1961), and will now be stated.

An IC grammar cannot differentiate governors from dependents, but it can differentiate degrees of closeness of relation. Thus a dependency grammar for English might treat verbs as governors of their subjects and objects, whereas an IC grammar might treat verb and object as partners in a constituent, the latter in turn being one partner in a subject-predicate construction. Since the two theories emphasize different aspects of linguistic patterning, it is reasonable to adopt a definition of structural correspondence that allows several dependency structures to correspond to one IC structure, and vice versa.

A SUBTREE is a connected subset of a tree. A complete subtree consists of some element of a tree, plus all others connected to it, directly or indirectly, and more remote from the origin of the tree. The tree in Figure 1 includes four complete subtrees: (X_{11}), (X_2, X_{32}), (X_{12}), and all four elements (since the tree is a subtree of itself).

An IC structure and a dependency structure, both defined over the same string, correspond relationally if every constituent is coextensive with a subtree and every complete subtree is coextensive with a constituent. (Two structural entities are coextensive if they refer to the same elements of a terminal string.) The sense of relational correspondence is plainly to fix requirements for similarity of unlabeled structural diagrams; this is the extent of Hays's definition.

Gaifman adds a restriction on labeling; its necessity is apparent from Figure 2, where an ambiguous sentence is shown with two IC diagrams and two dependency trees. Structures (a) and (c) take *are flying* as a compound verb; structures (b) and (d) take *flying planes* as a noun phrase. Yet

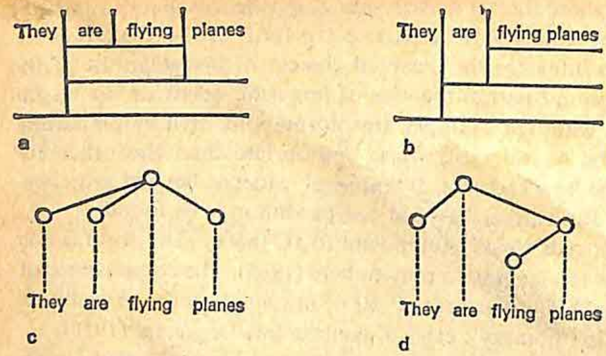


Figure 2 IC and dependency structures for an ambiguous sentence

structure (c) corresponds relationally to both (a) and (b). The requirement of label correspondence avoids such counterintuitive matches.

In dependency theory, a string DERIVES FROM the auxiliary that labels the root node in a corresponding tree. Thus every string in the set characterized by a dependency grammar derives from one of the auxiliaries X_i such that $*(X_i)$ is a rule in the grammar. In IC theory, a string derives from the auxiliary that labels it as a constituent. These definitions are valid for all strings covered by complete subtrees, on the one hand, or by constituents, on the other. An auxiliary accounts for the set of strings that derive from it; two auxiliaries are SUBSTANTIVELY EQUIVALENT if they account for the same set of strings.

An IC structure and a dependency structure CORRESPOND if they correspond relationally and every complete subtree has an origin labeled with an auxiliary that is substantively equivalent to the auxiliary labeling the coextensive constituent.

A dependency grammar is strongly equivalent to an IC grammar if they have the same terminal alphabet, and for every string over that alphabet, every structure attributed by either grammar corresponds to a structure attributed by the other. It follows that the grammars are weakly equivalent, since if one assigns no structure to a string the other must do likewise.

Gaifman shows that every dependency grammar is strongly equivalent to an IC grammar, but that IC grammars with a property he establishes are not strongly equivalent to any dependency grammars. Hence the two theories are not strongly equipotent. The property of some IC grammars that cannot be imitated in dependency theory cannot be stated in a few words; it seems most unlikely that linguistic applications would call for grammars with the property, and if that be true, dependency theory is in that respect preferable: the weakest empirically adequate theory is always best, other criteria failing to distinguish. However, the empirical point is unsettled.

As for the relative economy of dependency and IC theories, nothing is presently known. The intuitive appeals of the two theories cannot be discussed, since intuitions are personal and irrational. Other bases for comparison are adduced in the following sections.

7 Extensions of dependency theory

In contemporary linguistic theory, two major avenues for extension of syntactic models have been opened. One is transformation theory, which extends context-free IC grammars by admitting additional kinds of rules. The other is stratification theory, which extends IC grammars by combining them in sequences. Proponents of both systems claim greater power, economy and intuitive acceptability.

Both avenues do guarantee greater formal power. Dependency and IC theories, weakly equipotent to one another, can characterize only a subclass of the sets of strings characterizable by transformational or stratified systems. Economy, in a sense, can be improved by following either avenue, although discussions of this point generally fail to mention the loss of economy incurred by admitting more powerful rules or sequences of grammars. Economy, in any case, is often a matter of notational conventions, achievable without a change in power.

A transformation is a rule that replaces an ordered k-tuple of IC structures, or P-markers, with a single structure. One trite example is the combination of structures associated with *the old man* and *the tall man* to yield the structure associated with *the tall old man*. The argument for use of a transformation rule in this case is that an IC grammar must overspecify the structure by making *old man* a constituent of *tall old man*; otherwise, an IC rule would be needed for a noun with one adjective, another rule for a noun with two, another for three, and so on with no obvious limit. Other transformations are proposed for passive voice, for conjunction, and so on.

Rules that compose k-tuples of dependency trees into single trees can be defined in like manner. This proposal does not fall within the scope of Chomsky's strictures on transformation rules applied to k-tuples of strings, since dependency trees carry structure information (1961). Casual examination suggests that there would be little difference between transformation of dependency trees and transformation of IC structures, but no definite investigation has been undertaken.

In a stratified grammar, each stratum has its own alphabet, tactic rules and so on; Lamb postulates four strata, which he calls phonemic, morphemic, lexemic, and sememic (1964a). A string over the alphabet of any stratum has three kinds of properties. First, it does or does not belong to a set characterized by a grammar for its stratum. Secondly, it is related, by a transduction, to a single string or set of alternative strings on the adjacent higher stratum. And thirdly, it is similarly related to a string or set of strings on the adjacent lower stratum. (Exceptions must naturally be made for strings on the highest and lowest levels.)

Empirically, the set of strings on the lowest stratum is of primary interest, since these strings, if any, are testable for sentencehood. The set of principal interest is not merely the set characterized by the lowest stratal grammar (i.e. the phonemic), but the subset that maps, by transduction, onto a certain set of morphemic strings. Again, the set at issue is not the whole set characterized by a morphemic grammar, but a subset. In short, a terminal string must satisfy independent phonemic, morphemic, lexemic and sememic criteria.

The application of several grammars is tantamount to taking the intersection of several languages. Let us suppose, for example, that the grammars of the four strata are context-free, and that the transductions, or conversions, from stratum to stratum are finite-state. Then the conversions have no effect on the power of the system, since a finite-state transduction maps any context-free language onto a context-free language (Ginsburg-and Rose, 1963). The system thus characterizes the intersection of four context-free languages; the intersection, not necessarily context-free, is context-sensitive[2]. The result is the same even if only two of the grammars are context-free; those of the phonemic and morphemic strata could be taken as finite-state with no loss in power, and indeed both Hays (1961) and Lamb (1964b) have made such suggestions.

Transductions are not necessarily finite-state, although such operations as consultation of a dictionary are readily formalized with finite-state machines (Lamb and Jacobsen, 1961). Pushdown transducers are essentially more powerful; one possible application in stratification theory is in generative systems or decision procedures with vacuous grammars for all but one of the strata. A sequence of finite-state transductions up to the sememic stratum, followed by application of a context-free grammar, is a decision procedure weakly equipotent to direct application of a context-free grammar on the lowest stratum. (Questions of economy are not being considered here.) On the other hand, a sequence including at least one pushdown transduction and ending with application of a context-free grammar is more powerful.

A significant aspect of stratification theory is its demand for TRANSCRIPTION of the text on each stratum. If Lamb's four strata are adopted, four alphabets must be posited. An element of the alphabet on any stratum is said to be REPRESENTED by an element or combination of elements on the next lower stratum. On each stratum, moreover, the law of composition of alphabetic elements can be chosen freely. In linguistics, the tradition is to deal always with strings, associating structural descriptions (e.g. P-markers) with them. On the phonemic stratum, and even perhaps on the morphemic stratum in Lamb's model, this tradition need not be altered. On higher strata, other composition laws are attractive. Instead of concatenating lexemic elements into strings, we can concatenate them, by a different law, into trees. In particular, we can regard DEPENDENCY TREES not as structural descriptions of strings but as the fundamental transcriptions of utterances on their stratum. For the sememic stratum, a still more complex law is wanted; the transcription there takes the form of a GENERALIZED GRAPH, or network.

---

[2] Sheila A. Greibach, personal communication.

If this view is adopted, the transduction from the morphemic stratum to the lexemic can be regarded as having two stages. The first, to be understood intuitively as the consultation of a dictionary, passes from a string of morphemic elements to one of lexemic units. The second, very much like a decision procedure for a dependency grammar, passes from a lexemic string to a lexemic tree. A similar account could be given of transduction to the sememic stratum, but here trees must be mapped onto graphs.

If a dependency tree is to be the transcription of an utterance, it must be labeled at each node with sufficient information to identify the corresponding phonemic transcription, at least up to the point of free lexemic, morphemic and phonemic variation. Yet the amount of information required at each node is far less than that given by the corresponding element of a lexemic string; some of the information given by the identity AND ORDER of elements in the string is also given by the grammatical labels and dependency relations in the tree. To record only the string is to leave implicit dependency relations which, explicitly recorded, are helpful in conversion to the next-higher stratum; to record both tree and string introduces redundancy. To take the tree as lexemic transcription, with only the necessary supplementary labeling at each node, resolves the dilemma.

In fact, it is possible to construct a dependency grammar in such a fashion that the order of elements in lexemic strings, and all so-called GRAMMATICAL elements (such as affixes of agreement), are determined from the labels of nodes and the dependency relations during transduction downward.

A grammar composed in this manner eliminates the inefficiency, mentioned in section 2, of assigning a unit such as *book-* to two categories, N_{sg} and N_{pl}. The double assignment was made only for the sake of agreement with verbs. According to the new plan, only the subject (or only the verb) would have fixed grammatical number in the lexemic tree transcription. The relation of subject to verb would be differentiated from that of object to verb. In the transduction to the morphemic stratum, the order of these elements would be determined and the number of the unmarked element would be decided in accordance with the agreement rule, now a rule of transduction.

8 Empirical questions

The distinction between formal models and classes of empirical data, or procedures for data collection, is sometimes lost in linguistic debates. If the data of linguistics are restricted to observations of GRAMMATICALITY or SENTENCEHOOD, then no empirical reason for preferring the dependency model to that of phrase structure – or vice versa – can be given, since the two are weakly equipotent. If instead the range of data to be explained by linguistic theory is widened, empirical grounds for choice can perhaps be found. This statement is still quite far removed from any statement associating dependency theory with a specific interpretation.

One line of interpretation would make dependency a semantic theory, justifying the valences in any grammar by reference to meaningful relations among elements. Naturally, in order to give this interpretation an empirical basis, techniques for collecting data about meaningful relations would have to be devised. As Garvin has pointed out, translation and paraphrase give at least indirect evidence about meaning; the task of obtaining semantic data is not hopeless (1958).

As an argument favoring adoption of a dependency model, this one is potentially interesting. It can be put in terms of simplifying the transduction between two strata (Lamb's lexemic and sememic). It provides a rationale for counting co-occurrences of elements. If phrase-structure theory is used, the only pairs of elements that can be said to co-occur in text are those that occur as partners in simple constructions. When one of the partners is complex, the elementary partner co-occurs with the complex, not with any of its elementary constituents. In the dependency structure of a sentence, every element co-occurs with its governor and with its dependents, if any.

Another line of interpretation makes dependency a psychological relation. Empirically, this interpretation would require data about the structural intuitions of native speakers; dependency theory would be supported if many or most native speakers could agree on the central (governing) element in each utterance presented to them, and on the connections binding elements together. Phrase-structure theory would be supported if they agreed on containments, e.g. that object-verb relations are closer than subject-verb relations, so that predicates are contained in sentences.

Psycholinguistic material is notoriously hard to manage. Naïve native speakers of Standard Average European are almost extinct; readily obtainable data are bound to be contaminated with school-learned rules that in no way affect the speaker's language system outside the observation room. The linguist need not conclude, however, that psycholinguistic data are totally forbidden to him, but only that their use requires proper controls.

Like the semantic argument, the psychological justification for dependency theory is potentially interesting. Since neither kind of data is presently available, there is no reason to attach dependency theory to either. Recent formal work in linguistics has drawn attention to the inadequacy of sentencehood data; at best, use of such data must leave the linguist with a plurality of theoretical models, each capable of explaining his facts. Turning to other varieties of data seems, at this time, to offer the best hope of deciding relatively delicate issues like the choice between IC and dependency theories, or between transformation and stratification.[3]

---

[3] I am especially grateful to Charles F. Hockett, Martin J. Kay, Sydney M. Lamb and Yves Lecerf for discussions of the subject of this paper.

References

BAR-HILLEL, Y., PERLES, M., and SHAMIR, E. (1961), 'On formal properties of simple phrase structure grammars', *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*, vol. 14, pp. 143–72.

BERGE, C. (1959), *Théories des Graphes et ses Applications*, Paris.

CHOMSKY, N. (1956), 'Three models for the description of languages', *IRE Trans. PGIT*, vol. 2, pp. 113–24.

CHOMSKY, N. (1961), 'On the notion "rule of grammar"', *Proc. Symposia appl. Math.* vol. 12, p. 18.

CHOMSKY, N. (1963), 'Formal properties of grammars', in R. Duncan Luce, Robert R. Bush and Eugene Galanter (eds.), *Handbook of Mathematical Psychology*, Wiley.

CHOMSKY, N., and MILLER, G. A. (1963), 'Introduction to the formal analysis of natural languages' in R. Duncan Luce, Robert R. Bush and Eugene Galanter (eds.), *Handbook of Mathematical Psychology*, Wiley.

GAIFMAN, H. (1965), 'Dependency systems and phrase structure systems', *Information and Control*, vol. 8, pp. 304–37.

GARVIN, P. L. (1958), 'A descriptive technique for the treatment of meaning', *Language*, vol. 34, pp. 1–32.

GINSBURG, S., and ROSE, G. F. (1963), 'Operations which preserve definability in languages', *JACM*, vol. 10, pp. 175–95.

GORN, S. (1961), 'Specification languages for mechanical languages and their processors – a baker's dozen', *Communc. ACM*, vol. 4, pp. 535–7.

HAMP, E. P. (1963), 'The transformation theory', *Georgetown University Monograph Series on Language and Linguistics*, no. 15, pp. 3–38.

HARRIS, Z. S. (1946), 'From morpheme to utterance', *Language*, vol. 22, pp. 161–83.

HAYS, D. J. (1961), 'Linguistic methodology and the theory of strata', paper presented at a meeting of the Philological Association of the Pacific Coast, November.

HAYS, D. J. (1961), 'Basic principle and technical variations in sentence-structure determination', in Colin Cherry (ed.), *Information Theory*, pp. 367–76. (*Proceedings 4th London Symposium Information Theory*, Royal Institution, 1960.)

HAYS, D. J. (1961), 'Grouping and dependency theories', in H. P. Edmundson (ed.) *Proceedings National Symposium on Machine Translation*, Prentice-Hall, pp. 258–66.

HAYS, D. J. (1962), 'Automatic language-data processing', in Harold Borko (ed.) *Computer Applications in the Behavioral Sciences*, Prentice-Hall, pp. 394–421.

HJELMSLEV, L. (1961), *Prolegomena to a Theory of Language*, trs. Francis J. Whitfield, Madison.

IHM, P., and LECERF, Y. (1963), *Eléments pour une grammaire générale des langues projectives*, Brussels.

IORDANSKAJA, L. N. (1963), 'O nekotoryx svojstvax pravil'noj sintaksičeskoj struktury', *Voprosy jazykoznanija*, vol. 4, pp. 102–12.

LAMB, S. M. (1964a), 'The sememic approach to structural semantics', in Roy d'Andrade and A. Kimball Romney (eds.), *Transcultural Studies in Cognition*, *Amer. Anthropol.*, vol. 66, no. 3, pt 2, pp. 57–78.

LAMB, S. M. (1964b), 'Stratificational linguistics as a basis for mechanical translation', Unpublished paper presented at the US–Japan Seminar on Mechanical Translation.

LAMB, S. M., and JACOBSEN, W. H. Jr (1961), 'A high-speed, large-capacity dictionary system', *Mechanical Translation*, vol. 6, pp. 76–107.

LECERF, Y. (1960), 'Programme des conflits, modèle des conflits', *La Traduction Automatique*, vol. 1, no. 4, pp. 11–20, no. 5, pp. 17–36.

17 R. H. Robins

Basic Sentence Structures in Sundanese

R. H. Robins, 'Basic sentence structures in Sundanese', *Lingua*, vol. 21, 19 8, pp. 351–8.

It is hoped that this short article may constitute some slight contribution to our knowledge of the Sundanese language and to the formulation of linguistic descriptions. It is therefore a pleasure to offer it, for what it may be worth, to the honouring of a distinguished linguistic scholar in the Netherlands, a country to whose scholarship we owe so much both for our understanding of the languages of Indonesia and for progress in linguistic science.

An attempt is here made to set out the major underlying sentence structures of the Sundanese language and the categories that are applicable to them. Some of these structures may correspond in part to the output of the phrase-structure component of some types of generative grammar, and others may be transformationally derivable from them; but these questions are not pursued. The sentence structures are describable as basic, in that all the elements in them are either obligatory, or else are involved in a specific relationship with the total structure and are thus peculiar to one or more basic structures and not possible members of all structures. Thus the 'indirect object', *ka* N(oun), is peculiar to sentences containing a ditransitive verb (see p. 243) and is therein part of the basic structure of such sentences, even though it may be optionally omitted (or in generative terms 'deleted').

These basic sentence structures form the nuclei of large numbers of expanded structures comprising the favourite sentence types of the language. Expansions that are in principle available to specific elements in all the basic structures and are therefore characteristic of none of them are not discussed here: e.g. particles, *gɤs*, already, *ayɤna*, now; auxiliary verbs, *kudu*, should; adverbial phrases, *di kota*, in the town: adjectives, (*kota*) *gəde*, large (town); subordinate clauses, *waktu* ..., when, *supaya* ..., in order that, *nu* ..., who, which; subordinating nominalization, *bapa dataŋna*, father's arrival (*bapa dataŋ*, father arrives). Nor are expansions discussed that are restricted to certain lexical sets of words: e.g. *ŋaʤalanan ka Banduŋ*, to make a journey to Bandung, but presumably not **sare ka Banduŋ*, *to sleep to Bandung; *ŋomoŋ yen* ..., to say that ..., but presumably not **ŋaʤalanan yen* ..., *to make a journey that .... Also excluded from attention are constructions involving words whose syntax is in certain respects individual to the word concerned; e.g. *ʤadi*, to be, to become; *aya*, there is. The construction types that are considered here are favourite sentence types in that a large and often open-ended set of words may appear in the places occupied by the various elements and their expansions.

It is felt that the sentence structures here treated may be seen to underlie the great majority of the sentences in actual discourse, and that the material excluded from treatment here would best be dealt with on the basis of the structures here set up without requiring major alterations in the descriptions that are given. It is further considered that more delicate subclassifications of the word classes involved as elements of structures will not invalidate these basic structures but refine them.

The syntactically relevant word classes in this study are N(ouns), V(erbs), and Particles; adjectives may be regarded as a subclass of (intransitive) verbs, and pronouns as a subclass of nouns; words that it is sometimes convenient to designate as adverbs may be assigned in part to the class of particles and in part to a syntactically restricted subclass of nouns.[1]

Sundanese sentences may be equational or verbal.

Equational sentences consist basically of two noun elements (or, by expansion, of two nominal phrases or nominal groups); the first noun element, in unemphatic or unmarked sentences, may be a pronoun.

Exx. *manehna praʤurit*, he (is) a soldier
patlot iɤ pamere, this pencil (was) a present.

This structural type, though productive, is less frequent than that of the verbal sentences.

Verbal sentences consist basically of a noun element followed, in unmarked sentences, by a verb element.

Exx. *manehna sare*, he is sleeping
kapal iɤ dataŋ, this ship is arriving.

In both equational and verbal sentences the (first) noun element may, grammatically, be omitted; but except in imperative sentences (see p. 244) this is unlikely outside context-bound sentences in continuous discourse.

Verbal sentences may be subdivided according to the syntactic subclassification of the verb.

Intransitive verbs enter into basic structures containing one noun element: N V, *manehna sare*, he is sleeping.

---

[1] W. B. Van Syoc (1959), while formally recognizing the grammatical kinship between the Sundanese adjective and verb, sets out diagnostic frames for the recognition of the four classes: verb, noun, adjective, and particle.

Transitive verbs enter into basic structures containing two noun elements: N_1 V N_2, *murid mawa buku*, the pupil carries the book; *kuriŋ ŋarekɤn budak*, I send the child to sleep.

Ditransitive verbs enter into basic structures containing three noun elements: N_1 V N_2 *ka* N_3, *guru mere buku ka murid*, the teacher gives the book to the pupil. N_3 is preceded by *ka*, to, but in this construction *ka* N_3 is part of the basic structure, and not simply an adverbial phrase like *ka Banduŋ* in the example above (p. 241), because N_3 enters into the corresponding passive sentences as a specific structural component (see p. 245).

In the examples already given, and those to be given later, the word order indicated is the normal order for unmarked, non-emphatic, sentences; several variations of word order within the structures are possible, often accompanied by a specific intonation tune, emphasizing one or another element in the sentence.

The basic categories of the verb are active and passive, and they are distinguished formally in both syntax and morphology. These categories and the subdivisions of the verb class already mentioned apply both to simple (non-derived) verb forms and to the various classes of derived verbs that are a conspicuous part of Sundanese morphology.[2]

Intransitive verbs bear only the active category; transitive verbs bear both active and passive categories. The active category is characterized syntactically by the structures N V, N_1 V N_2, and N_1 V N_2 *ka* N_3; the passive category is characterized by the structures N_2 V *ku* (by) N_1, N_2 V *k*a N_3 *ku* N_1, and N_3 V N_2 *ku* N_1. In regularly formed transitive verbs the active category is characterized morphologically by the nasalized forms,[3] and the passive category is characterized by the non-nasalized forms with or without the prefix *di-*.

It may be noted that the morphological distinction between these two forms of the verb underlies a large number of deverbal derivational forms, both nouns and verbs.[4]

The sentence categories, declarative, interrogative and imperative are imposed on sentences containing one or other of the two verbal categories:

Declarative is indicated by statement intonation, usually involving a final fall, and with passive verbs by the presence of the prefix *di-* (or, less commonly, *ka-*).[5]

---

[2] The morphology of Sundanese verbs is dealt with in Van Syoc's thesis and in Robins (1959).

[3] For the details of the nasalization process, see W. J. Van der Dussen and D. J. Duijnhouwer (1950), pp. 29–31 and Robins (1953).

[4] See further the references cited in footnote 2.

[5] In older periods of the language it seems that the prefix was not an obligatory exponent of the declarative category in passive verbs (Eringa, 1949). This gives

Interrogative is indicated by a specifically interrogative intonation, involving a final rise (symbolized by *?*), and often additionally by a specific question word (*saha*, who?, *naon*, what?, *kumaha*, how?, etc.) or by the particle *teh*.

Imperative positive is indicated by a specific intonation (symbolized by *!*), and with passive verbs by the absence of the prefix *di-*. Imperative negative (prohibitive) is indicated by *ulah* preceding a declarative structure, usually accompanied by an imperative intonation.

The categories of declarative and interrogative apply also to equational sentences.

The following examples will illustrate the categories so far described:

Equational sentences

Declarative N N	*manehna praʤurit,* he is a soldier
Interrogative N N	*manehna teh praʤurit?,* is he a soldier? *naon ɤsina?*, what are the contents?

Verbal sentences

Intransitive	*active*
Declarative N V	*manehna lɤmpaŋ*, he is running *bapa dataŋ*, father is arriving, has arrived
Interrogative N V	*manehna teh lɤmpaŋ?*, is he running? *bapa teh dataŋ?*, is father arriving?, has father arrived?
Imperative (N) V	(*maneh*) *lɤmpaŋ!*, (you) run! (*maneh*) *ulah lɤmpaŋ!*, don't (you) run!

Transitive	*active*	*passive*
Declarative N_1 V N_2	*kuriŋ mɤli buku*, I buy a book	N_2 V *ku* N_1 *buku dibɤli ku kuriŋ*, a book is bought by me
Interrogative N_1 V N_2	*maneh teh mɤli buku?*, are you buying a book?	N_2 V *ku* N_1, *buku teh dibɤli ku maneh?*, is the book being bought by you? *buku dibɤli ku saha?* by whom is the book being bought?

---

historical reinforcement to the description of the contrast between nasalized active and non-nasalized passive as the basic formal distinction in Sundanese verbs.

Imperative (N_1) V N_2	*(maneh) mɤli buku!*, (you) buy a book!	V N_2 (*ku* N_1), *bɤli buku* (*ku maneh*)*!*, buy the book![6]
	(*maneh*) *ulah mɤli buku!*, don't (you) buy a book!	*ulah dibɤli buku* (*ku maneh*)*!* don't (you) buy the book!

Ditransitive verbs enter into two types of structure, depending on whether the *ka* preceding N_3 is suppressed with a corresponding change of normal word order.

	active	*passive*
Declarative N_1 V N_2 *ka* N_3	*guru mere buku ka murid*, the teacher gives a book to the pupil	N_2 V *ka* N_3 *ku* N_1, *buku dibere ka murid ku guru*, the book was given to the pupil by the teacher
N_1 V N_3 N_2	*guru mere murid buku*, the teacher gives the pupil a book	N_3 V N_2 *ku* N_1, *murid dibere buku ku guru*, the pupil was given the book by the teacher
Interrogative N_1 V N_2 *ka* N_3	*guru teh mere buku ka murid?*, does the teacher give the book to the pupil?, etc.	(the corresponding two passive structures need not be given)
Imperative (in the N_2 N_3 passive imperative *ka* may be optionally omitted)		
(N_1) V N_2 *ka* N_3	(*maneh*) *mere buku ka murid!* (you) give a book to the pupil!	V N_2 (*ka*) N_3 (*ku* N_1), *bere buku* (*ka*) *murid* (*ku maneh*)*!*, (you) give the book to the pupil!
(N_1) V N_3 N_2	(*maneh*) *mere murid buku!*, (you) give a book to the pupil!	V N_3 N_2 (*ku* N_1), *bere murid buku* (*ku maneh*)*!* (you) give the pupil a book!

(The corresponding imperative negative structures need not be given).

A second class of ditransitive verbs is formed by the affixation of *paŋ-* and *-kɤn* on to transitive verb forms; these forms have a class meaning of benefactive action: e.g. *bawa, mawa* to carry; (*di*)- *paŋbawakɤn, maŋ-*

---

[6] Van der Dussen and Duijnhouwer distinguish the active and passive imperatives as subjective and objective respectively, the former being used when the object noun is not emphasized, the latter when the object noun is emphasized or is made definite, e.g. by *eta*, that. But my informant made it clear that for him the passive imperative was the normal, unmarked, form.

mawakɤn, to carry (something) for someone else. With these verbs the preposition preceding N_3 is usually *kɤr*, for, on behalf of, but in some cases it is *ka*, to. This preposition may often be omitted, and then the usual word order is N_1 V N_3 N_2.

At least some of these verbs also enter into two types of passive structure, and their forms differ accordingly. When N_2 is preposed to the verb, N_3 retains *kɤr* (or *ka*), and both the prefix and the underlying verb appear in their non-nasalized forms (e.g. *(di)paŋbawakɤn*, to be carried for someone else). When *kɤr* (or *ka*) is obligatorily suppressed and N_3 is preposed to the verb, the underlying verb appears in its nasalized form (e.g. *(di)-paŋmawakɤn*, to have (something) carried for one). Both types of structure are available in passive imperative sentences.

	Active	*Passive*
Declarative N_1 V N_2 *kɤr* N_3	*murid maŋmawakɤn buku kɤr guru*, the pupil carries the book for the teacher	N_2 V *ku* N_1 *kɤr* N_3, *buku dipaŋbawakɤn ku murid kɤr guru*, the book is carried by the pupil for the teacher
N_1 V N_3 N_2	*murid maŋmawakɤn guru buku* the pupil carries the book for the teacher (*bapa maŋmaʧakɤn surat ka anakna, bapa maŋmaʧakɤn anakna surat*, father reads the letter to to his child)	(*surat dipaŋbaʧakɤn ku bapa ka anakna*, the letter is read by father to his child)
		N_3 V N_2 *ku* N_1, *guru dipaŋ mawakɤn buku ku murid*, the teacher has the book carried for him by the pupil (*anak dipaŋmaʧakɤn surat ku bapana*, the child has the letter read to him by his father)
Interrogative N_1 V N_2 *kɤr* N_3	*murid teh maŋmawakɤn buku kɤr guru*, does the pupil carry the book for the teacher?, etc.	(the corresponding passive structures need not be given)

Imperative		
(N_1) V N_2 *kɤr* N_3	(*maneh*) *maŋmawakɤn buku kɤr guru!*, (you) carry a book for the teacher!	V N_2 kɤr M_3 (*ku* N_1), *paŋbawakɤn buku kɤr guru* (*ku maneh*)*!*, (you) carry the book for the teacher!
(N_1) V N_3 N_2	(*maneh*) *maŋmawakɤn guru buku!*, (you) carry a book for the teacher!	V N_3 N_2 (*ku* N_1), *paŋmawakɤn guru buku* (*ku maneh*)*!*, (you) carry the book for the teacher!

(the corresponding imperative negative structures need not be given).

From these basic sentence structures one may summarize the nuclear classes and categories of Sundanese grammar as comprising:

Two sentence types: equational, verbal;
Three sentence categories: declarative, interrogative, imperative;
Three word classes: noun, verb, particle;
Three verbal subclasses: intransitive, transitive, ditransitive;
Two verbal categories: active, passive.

That these categories and classes and their interrelations show similarities at several places to partially corresponding categories and classes in other languages is not surprising. Equally, one must recognize in their particular configurations and distributions part of the individuality of the Sundanese language.

References

ERINGA, F. S. (1949), 'Loetoeng kasaroeng: een mythologisch verbaal uit West-Java', *Verhandelingen van het koninklijk Instituut voor taal-, land-en Volkenkunde*, pp. 85–91.

ROBINS, R. H. (1953), 'The phonology of the nasalized verbal forms in Sundanese', *BSOAS*, vol. 15, pp. 138–45.

ROBINS, R. H. (1959), 'Nominal and verbal derivation in Sundanese', *Lingua*, vol. 8, pp. 337–69.

VAN DER DUSSEN, W. J., and DUIJNHOUWER, P. J. (1950), *Sundaas voor de praktijk*, Mouton.

VAN SYOC, W. B. (1959), *The Phonology and Morphology of the Sundanese Language*, Ph.D. thesis, Ann Arbor.

18 M. A. K. Halliday

Options and Functions in the English Clause[1]

M. A. K. Halliday, 'Options and functions in the English clause',
Brno Studies in English, vol. 8, no, 137, 1969, pp. 81–8.

In his excellent summary of the work of the Prague school, Josef Vachek draws attention to the development by Czechoslovak linguists of the 'functionalist' view of linguistic structure (Vachek, 1966). One of the characteristics of this approach has been the recognition of several components in the organization of the grammar of a language, a conception which Vachek shows to be derivable from the work of Mathesius and of Bühler. The importance of this conception appears very clearly from a study of the systems and structures of the English clause.

The systems having the clause as their point of origin group themselves into three sets which I have referred to elsewhere under the headings of transitivity, mood and theme (Halliday, 1967b). These labels refer specifically to sets of clause systems, which are however relatable to these general components of the grammar. Those of transitivity belong to that area which Vachek derives from Bühler's 'Darstellungsfunktion' and glosses as 'informing of the factual, objective content of extralinguistic reality' (Vachek, 1966); Daneš calls it the 'semantic structure' of the sentence (1964)[2]. Those of mood express speech function, the relations among the participants in a speech situation and the speech roles assigned by the speaker to himself and his interlocutors; this includes most of Poldauf's 'third syntactical plan',[3] and embraces both of Bühler's additional functions – the speaker's attitude and his attempt to influence the hearer – though excluding (as outside grammar) the paralinguistic indexical signals. Theme is the clausal part of Mathesius's 'functional sentence analysis' (Vachek, 1966, pp. 59 ff.), Daneš' 'organization of utterance' (1964, pp. 227 ff.), which continues to be studied extensively by Firbas and others (1964; 1966); this concerns the structuring of the act of communication within the total framework of a discourse, the delimitation of message units and the distribution of information within them.

---

[1] My thanks are due to A. Henrici, R. D. Huddleston and R. A. Hudson, on whose ideas I have drawn freely in the preparation of this paper.

[2] Daneš (1964), 'based on . . . relations [which] are derived from nature and society . . . e.g. actor and action; the bearer of a quality or of a state and the state; action and an object resulting from the action or touched by it, etc.: different circumstantial determinations (of place, time etc.); causal and final relations; relations of consequence, etc.'. I would prefer to separate the 'cognitive' (factual, experiential) from the logical (see below).

[3] Poldauf (1964). But not, I think, the English beneficiary, which is a 'cognitive' function – though the very restricted 'ethic dative' does perhaps fall within this category.

Thus the English clause embodies options of three kinds, experiential, interpersonal and intratextual, specifying relations among (respectively) elements of the speaker's experience, participants defined by roles in the speech situation, and parts of the discourse. Although the clause options do not exhaust the expression of these semantic relations – other syntactic resources are available, quite apart from the selection of lexical items – the clause provides the domain for many of the principal options associated with these three components. At the same time it is useful to recognize a fourth component, the logical, concerned with the and's and or's and if's of language; this is often subsumed under the first of those above (e.g. by Daneš; see footnote 2) with some general label such as 'cognitive', but it is represented by a specific set of structural resources (hence not figuring among the clause options) and should perhaps rather be considered separately. Let us then suggest four such generalized components in the organization of the grammar of a language, and refer to them as the components of extralinguistic experience, of speech function, of discourse organization and of logical structure.

The first three then enter into, and collectively exhaust, the determination of English clause structure. In other words, structural function in the clause is fully derivable from systems of options in transitivity, mood and theme. But no one of these sets of options by itself fully specifies the clause structure; each one determines a different set of structural functions. Deriving from options in transitivity are functions such as actor, goal and beneficiary; from modal options, those such as subject, predicator and WH- element; from thematic options, functions such as theme, given and new. The same item occupies simultaneously a number of distinct 'roles' in the structure, so that the element of structure is a conflation of functions from different sources: in *John threw the ball*, *John* is at once actor, subject and theme.

Let us consider, as an example, a clause taken from a recording of spontaneous informal conversation, the final clause (beginning at *well then*) in the passage *and he was accepted, became a court painter, well then surely he must have made a fair amount of money*. Phonologically the clause has two tone groups, both tone 1, the first tonic (nucleus) at *surely* and the second at *money*:[4]

---

[4] For a discussion of intonation in relation to the grammar of the English clause see Halliday (1967c).

//1 well then surely// 1 he must have made a fair amount of money//

Structurally, the transitivity functions are actor *he*, process (*must*) *have made*, range *a fair amount of money*; modal (a) subject *he*, predicator *must have made*, complement *a fair amount of money*, (b) modality *surely . . . must* (*have made*); thematic (a) theme *well then+surely+he*, rheme *must have made a fair amount of money*, (b) conjunctive *well then*, anaphoric *he*, (c) given *well then+he must have made*, new *surely+a fair amount of money* (see Halliday, 1967a; 1967b). The structural description of the clause will be a string (since we are not considering group structure, and there is no embedding) of elements each of which is a complex of structural functions:

theme/conjunctive+modality/theme/new+actor/subject/theme/
anaphoric+process/predicator/modality+range/new

(omitting complement, rheme and given as not needing to be specified); together with the phonological specification of a tone group boundary after *surely* and of both tone groups as tone 1.

The structural functions, and the intonation features, can be shown to be the realizations of options in the three areas of transitivity, mood and theme. Thus, for example, the presence of the function 'subject' in the clause realizes the option 'indicative' in the mood system: that of 'actor' realizes the option 'extensive' in the transitivity system, and so on. The grammar of the clause may be represented as a network of such options, to each of which is attached a realization statement showing the structural 'output' of that option. In some cases, such as those just exemplified, the realization statement merely specifies the presence of a certain structural function; but in other cases it may have the effect of ordering one function with respect to another or to initial or final position in the syntagm. But it may also specify the conflation of two functions into a single element of structure: for example the option 'operative' conflates the functions of subject and actor.

Underlying the structural description of the clause, therefore, is a systemic description in which the clause is represented in terms of the options selected. Figure 1 shows a simplified systemic description of the clause cited above, while Figure 2 shows the portion of the clause system network from which that description is derived.[5] Figure 1 may be expressed lineally as follows (where / indicates simultaneity,: hierarchy, and I_1 denotes an intersection):

(major: (extensive: ((descriptive: + range)/middle))/(indicative: ((declara-

---

[5] For conventions see Halliday (1967a; 1967b; 1967c). For a systemic description of mood in French see Huddleston and Uren (1969).

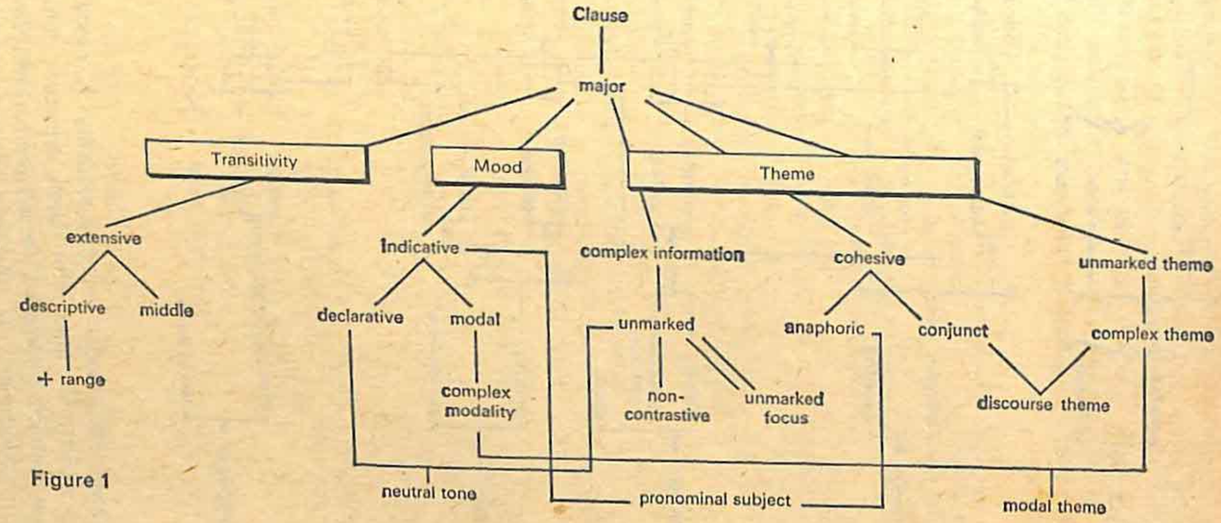

Clause
major
Transitivity
Mood
Theme
extensive
descriptive
middle
+ range
Indicative
declarative
modal
complex modality
complex information
unmarked
non-contrastive
unmarked focus
cohesive
anaphoric
conjunct
discourse theme
unmarked theme
complex theme
neutral tone
pronominal subject
modal theme

Figure 1

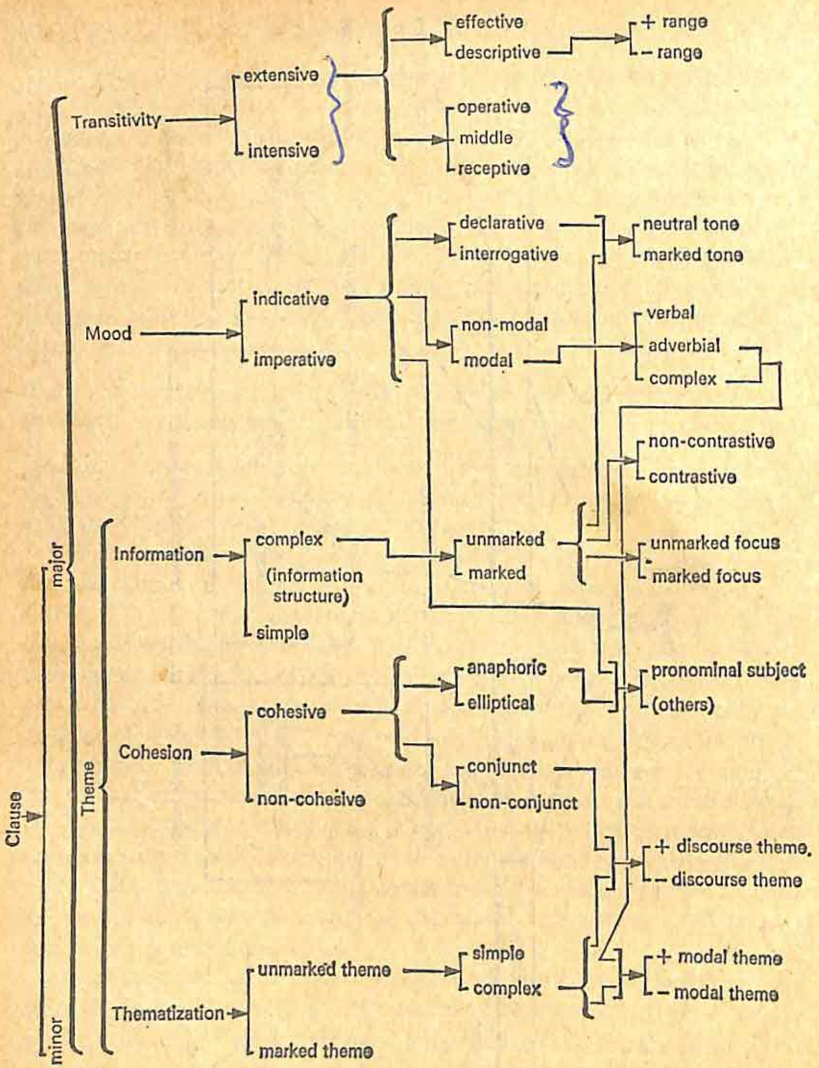


Figure 2

tive: I_1, neutral tone)/(modal: (complex modality: I_2, modal theme))/I_3, pronominal subject))/(complex information structure: unmarked information: (non-contrastive/unmarked focus/I_1/unmarked focus))/(cohesive: ((anaphoric: I_3)/(conjunct: I_4, discourse theme)))/(unmarked theme: complex theme: (I_4/I_2)))

We define a system as a set of options together with an entry condition, such that if the entry condition is satisfied one option from the set must be selected. The place of each option in the grammar is specified by the location, in the system network, of the system of which it is a term: thus the option 'declarative' contrasts with 'interrogative' in a two-option system the entry condition to which is 'indicative'. Table 1 lists, following

Table 1

Entry condition	*Option selected*	*Realization statement*	*Resultant partial structure*
Clause	major	insert Predicator/Process & Theme; Theme initial	*^Th·Pred/Proc·*
Major	extensive	insert Actor	*^Th·Pred/Proc·Ac·*
Major	indicative	insert Subject	*^Th·Pred/Proc·Ac·Su·*
Major	complex information structure	insert tone group boundaries initially & finally, plus one unplaced	*^Th·Pred/Proc·Ac· Su· //· //*
Major	cohesive	(none)	
Major	unmarked theme	conflate Subject with Theme	*//^Th/Su·Pred/Proc· Ac· //· //*
Extensive	descriptive	(specifies class of verb)	
Extensive	middle	conflate Actor with subject; (specifies voice of verb)	*//^Th/Su/Ac·Pred/ Proc· //· //*
Descriptive	+range	insert Range; concatenate with Process	*//^Th/Su/Ac·Pred/ Proc^Ra· //· //*
Indicative	declarative	concatenate Predicator with Subject	*//^Th/Su/Ac^Pred/ Proc^Ra· //· //*
Indicative	modal	insert Modality	*//^Th/Su/Ac^Pred/ Proc^Ra·Mo· //· //*
Modal	complex	insert Modality$_1$; conflate with Predicator	*//^Th/Su/Ac^Pred/ Proc/Mo*$_1$*^Ra·Mo· //· //*
Complex information structure	unmarked (information)	(specifies location of tone group boundary; see 'complex theme')	
Cohesive	anaphoric	insert Anaphoric	*//^Th/Su/Ac^Pred/ Proc/Mo*$_1$*^Ra·Mo· An· //· //*
Cohesive	conjunct	insert Conjunctive	*//^Th/Su/Ac^Pred/ Proc/Mo*$_1$*^Ra·Mo·An· Co· //· //*
Unmarked theme	complex theme	insert Theme$_1$ initially; (concatenate tone group boundary with it)	*//^Th*$_1$*^ //^Th/Su/Ac^ Pred/Proc/Mo*$_1$*^Ra· Mo·An·Co· //*

Table 1 – continued

Entry condition	*Option selected*	*Realization statement*	*Resultant partial structure*
Unmarked (information)	non-contrastive	insert (tone) 1 in first tone group	$//1^\wedge Th_1{}^\wedge//^\wedge Th/Su/Ac^\wedge Pred/Proc/Mo_1{}^\wedge Ra\cdot Mo\cdot An\cdot Co\cdot\ //$
Unmarked (information)	unmarked focus	insert New in each tone group; (conflate with final element)	$//1^\wedge Th_1\cdot Ne^\wedge//^\wedge Th/Su/Ac^\wedge Pred/Proc/Mo_1{}^\wedge Ra\cdot Mo\cdot An\cdot Co\cdot Ne^\wedge\ //$
Unmarked (information) & declarative	neutral tone	insert (tone) 1 in second tone group	$//1^\wedge Th_1\cdot Ne^\wedge\ //1^\wedge Th/Su/Ac^\wedge Pred/Proc/Mo_1{}^\wedge Ra\cdot Mo\cdot An\cdot Co\cdot Ne^\wedge\ //$
Anaphoric & indicative	pronominal subject	conflate Anaphoric with Subject	$//1^\wedge Th_1\cdot Ne^\wedge\ //1^\wedge Th/Su/Ac/An^\wedge Pred/Proc/Mo_1{}^\wedge Ra\cdot Mo\cdot Co\cdot Ne^\wedge\ //$
Conjunct & complex theme	discourse theme	conflate Conjunctive with Theme$_1$	$//1^\wedge Th_1/Co\cdot Ne^\wedge\ //1^\wedge Th/Su/Ac/An^\wedge Pred/Proc/Mo_1{}^\wedge Ra\cdot Mo\cdot Ne^\wedge\ //$
Complex modality & complex theme	modal theme	insert Theme$_2$; concatenate with Theme$_1$; conflate Modality with it	$//1^\wedge Th_1/Co^\wedge Th_2/Mo/Ne^\wedge\ //1^\wedge Th/Su/Ac/An^\wedge Pred/Proc/Mo_1{}^\wedge Ra/Ne^\wedge\ //$

the entry condition, each of the options appearing in the systemic description of the clause just cited, together with its realization statement; the final column shows the partial structure that results from the application of the realization statement. Since simultaneous options are unordered, the realizations could apply in a large number of different sequences; the partial structures shown here assume the options to be realized in the order listed. This does not affect the final 'partial structure', which is the structural representation of the clause and has the same form through whichever sequence of options it has been built up.[6]

Thus the clause is represented in the grammar as a set of features partially ordered in 'delicacy'; the features are options in the various systems of transitivity, mood and theme. The description shows not

---

[6] That is to say, indicative must be realized before declarative, and extensive before descriptive; but extensive and indicative are unordered with respect to each other and either may be realized first. The realization statements are given in a simplified form, appropriate only to the given environment and the given order of realization; moreover the systemic and structural descriptions are themselves simplified, a number of distinct options being sometimes conflated (so that the realization statement is a composite). Boundary symbols are omitted from the partial structures.

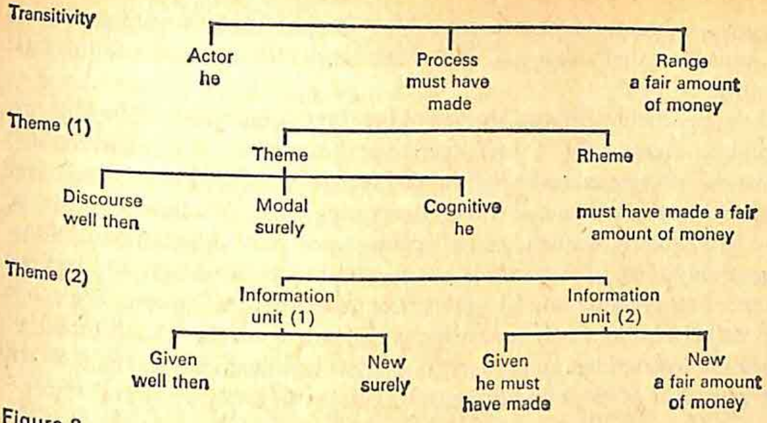


Figure 3

merely what options have been selected but also how these options are interrelated. The systemic representation is then realized as structure, with linear ordering; the example shows a simple string, though the total structural description at all ranks, and including any embedding, is a minimally bracketed tree structure (Hudson, 1967). Each element of this structure is a complex of functions, a set of structural 'roles' specified as realizations of the options selected. Options in transitivity, mood and theme all contribute to the specification of structural function. Not every clause constituent occupies a role in respect of all three – a modal adjunct, for example, has no transitivity role; on the other hand a constituent may have more than one role specified by one component, for example 'theme' and 'conjunctive' (both thematic functions), 'subject' and 'WH- element' (both modal functions). Moreover a role may extend over more than one element, for example rheme over process and goal, or modality over theme and rheme. In other words, the clause has a number of different but simultaneous constituent structures according to which set of options is being considered, as exemplified in Figure 3; the consequent overlapping of functions is shown in Table 2. This explains the minimal bracketing: such a

Table 2

Transitivity			| Actor	| Process	| Range
Mood (1)			| Subject	| Predicator	| Complement
Mood (2)		| Modal-	(	)-ity	|
Theme (1)	Theme$_1$	| Theme$_2$	| Theme$_3$	| Rheme	
Theme (2)	Given	| New	| Given		| New
	well then	surely	he	must have made	a fair amount of money

componential analysis of structural function presupposes a 'rank' (string constituent) rather than an 'immediate constituent' view of linguistic structure.

The assignment of clause options to the three components of transitivity, mood and theme reflects their interdependence: there is a relatively high degree of interdependence within each component and a relatively low degree (though not none) between the components. As already noted, the three are broadly defined. Mood includes not only the contextual functions of statement, question etc., but also options relating both to speaker's attitude and comment and to the structure of the predication. The range of possible attitudes and comments is dependent on the speech function: hence the restrictions on comment adjuncts in interrogative, the different interpretation of the same intonation pattern in declarative and imperative, and so on; likewise the 'key' of an assertion or command, the specificity of a question, the degree of commitment, reservation or certainty are all modal options (Halliday, 1967c). At the same time the predication is the carrier of speech function; it is the form of the predication that represents statement, question etc., and hence functions such as subject and predicator are modal functions. The term 'mood' refers to a set of related options which give structure to the speech situation and define the relations between speaker and interlocutors in a linguistic interaction.

Likewise theme subsumes not only theme – rheme and given – new but also other functions including those in various equative and substitutive structures, anaphoric reference etc.; the term 'functional sentence perspective' is itself a recognition of the interrelatedness of this area of clause organization. In many cases this is characterized by a marked/unmarked relation between systems; for example the association of theme with given is an unmarked one.[7]

How the components may be interrelated is seen for example in the active/passive distinction in English, which can best be understood as the interaction of thematic options with those of transitivity. There may also be overlap in realization: *can may might should* etc. have two distinct meanings, one experiential, related to *is able to* etc., the other modal, related to adjectives such as *possible*, only the latter having the full non-finite tense range. Similarly intonation figures as the realization of both modal and thematic options.[8] More generally, one set of options may act as an environment restricting the selection or determining the interpretation of options within the other: for example, the range of choices of theme is less restricted in declarative than in interrogative, and the interpretation of the equative as a thematic option depends on transitivity.[9] Interrelations of this kind link what are otherwise independent networks of systems. The grammar of the English clause may, it is suggested, be represented in three such networks, which specify the range of possible options and their conditions of entry. With an analysis of structural function into components, the structure of the clause is seen to be determined by the selections made by the speaker from among the options available.

---

[7] Firbas (1964) notes both the distinction and the relation between the two variables of given – new and theme – rheme.

[8] The modal being in general realized by pitch-glide features ('tone') and the thematic by features of delimitation and prominence ('tonality' and 'tonicity'), although the distinction is not absolute: certain tone selections are thematic rather than modal (e.g. the choice between tone 1 and tone 4 in declarative clauses shows both the

---

speaker's attitude, as being with or without reservation, and at the same time the relation of the clause to others in the discourse, tone 4 meaning 'but' and tone 1 meaning 'and'). It may be suggested that it is perhaps the very great part played by intonation in realizing thematic choices in English that has led to the suggestion, rightly rejected by Firbas (1966), that 'English is less susceptible to FSP than Czech'. It is difficult to account for the syntax of the English clause without regard to thematic organization.

[9] See the discussion in Firbas (1966) of the association of thematic structure with definiteness.

References

DANEŠ, F. (1964), 'A three-level approach to syntax', *Travaux Linguistiques de Prague*, vol. 1, pp. 225–40.

FIRBAS, J. (1964), 'On defining the theme of functional sentence analysis', *Travaux Linguistiques de Prague*, vol. 1, pp. 267–80.

FIRBAS, J. (1966), 'Non-thematic subjects in contemporary English', *Travaux Linguistiques de Prague*, vol. 2, pp. 239–56.

HALLIDAY, M. A. K. (1967a), 'Notes on transitivity and theme in English: I' *J. Linguistics*, vol. 3, no. 1, pp. 37–81.

HALLIDAY, M. A. K. (1967b), 'Notes on transitivity and theme in English: II', *J. Linguistics*, vol. 3, no. 2, pp. 199–244.

HALLIDAY, M. A. K. (1967c), *Intonation and Grammar in British English*, Mouton.

HUDDLESTON, R., and UREN, O. (1969), 'Declarative, interrogative and imperative in French', *Lingua*, vol. 22, pp. 1–260.

HUDSON, R. A. (1967), 'Constituency in a systemic description of the English clause', *Lingua*, vol. 18, no. 3, pp. 225–50.

POLDAUF, I. (1964), 'The third syntactical plan', *Travaux Linguistiques de Prague*, vol. 1, pp. 241–55.

VACHEK, J. (1966), *The Linguistic School of Prague*, Bloomington & London.

19 P. A. Reich

The Finiteness of Natural Language

P A. Reich, 'The finiteness of natural language', *Language*, vol. 45, 1969, pp. 831–43.

In 1956 Chomsky published a supposed proof that natural language cannot be produced by a finite state machine. This proof has been used by transformationally trained psychologists in their criticism of models used by other psychologists (e.g. Bever, Fodor, and Garrett, 1968); and it has been used by transformational linguists in their criticism of stratificational theory, particularly of the finite state representation of the theory implemented on the computer (Reich 1970a). I wish to show that this proof is invalid.[1]

Chomsky's proof is in two parts. The first part is a mathematical demonstration that a grammar which produces only mirror-image strings cannot be produced by a finite state device: I have no objection to this statement. The second part of the proof consists of the assertion that English contains such mirror-image constructions. The examples Chomsky gives are of two types: nested dependencies and centrally embedded relative clauses. Nested dependencies are sentences such as 1 and 2; central embeddings are sentences such as 3:

1. If S_1, then S_2.
2. Either S_3 or S_4.
3. The rat that S_5 ate the malt.

Chomsky claims that the constructions represented by the symbols S_1 through S_5 may be any English sentence, including sentences which themselves consist of nested dependencies and centrally embedded relative clauses. If we accept this last statement as true, then it is indeed true that English is a language which cannot be produced by a finite state device. But is it true? Consider these examples:

---

[1] This is a revision of a paper presented at the December 1968, LSA meeting. For their comments on earlier versions, I am indebted to James McCawley, Bruce Fraser, Paul Postal and John Ross, none of whom agree with the point of view expressed here. I am particularly indebted to Floyd Lounsbury for working through the Mohawk material with me.

4. The rat ate the malt.
5. The rat that the cat killed ate the malt.
6. *The rat that the cat that the dog worried killed ate the malt.
7. *The rat that the cat that the dog that the cow tossed worried killed ate the malt.

These sentences are all considered grammatical by Chomsky. I accept 4 and 5, but reject 6 and 7. In order to understand why Chomsky considers 6 and 7 grammatical, let us explore his definition of grammaticality. He originally stated that 'Every grammar will have to meet certain external conditions of adequacy', and gave the specific criterion that 'the sentences generated will have to be acceptable to the native speaker' (1957, p. 48). He even gave a test sufficient to claim ungrammaticality: if a native speaker reads a sentence with the same intonation given a list of unrelated words, then that sentence is ungrammatical (1957, p. 16). I agree with this kind of definition of ungrammaticality. It is clear and reasonable; more important, it is testable. In fact such tests have been performed by Miller (1962), who reports that subjects repeated sentences such as 6 and 7 with an intonation pattern that was 'characteristic of the recitation of a list of unrelated phrases'. More recently Marks (1968) has asked subjects to grade sentences as to grammaticality. He reports that 'one embedding was almost never judged to be ungrammatical, while two embeddings were always considered ungrammatical.'

Here we see a discrepancy between what Chomsky called grammatical and, opposed to this, tests of native English speakers which show that they do not accept sentences which have been centrally embedded more than once. Chomsky's response (1965, p. 11) was:

> The notion ACCEPTABLE is not to be confused with GRAMMATICAL. Acceptability is a concept that belongs to the study of performance, whereas grammaticalness belongs to the study of competence. . . . Grammaticalness is only one of many factors that interact to determine acceptability. Correspondingly, although one might propose various operational tests for acceptability, it is unlikely that a necessary and sufficient operational criterion might be invented for . . . grammaticalness.

Thus, when confronted by adverse data, Chomsky retreated from his empirical position of 1957, to a theory that he himself admits cannot be tested empirically. What is the reason for this drastic retreat from the data? Why did Chomsky not simply revise his grammar so that it agreed with the data? The answer appears in this passage (1965, p. 11–12):

> 'Note that it would be quite impossible to characterize the unacceptable sentences in grammatical terms. We cannot formulate particular rules of grammar in such a way as to exclude them.'

Thus we see that the reason for his not taking this obviously superior course is that he couldn't see how to do it.

Let us return to Chomsky's proof. Are sentences like 1 and 2 acceptable when embedded more than once?[2] Consider the following sentences:

8. ?If if my knight takes your pawn, then your bishop would take my knight, then I should move my king.
9. ?Either either Archibald will marry Adelaide, or Adelaide will elope with Rodney, or the sponsor will cancel the soap opera.

I find these sentences questionable at best. But even if one accepts them, an additional embedding, as in the following examples, is clearly unacceptable:

10. *If if if my knight takes your pawn, then your bishop would take my knight, then my queen would take your pawn, then you should move your king.
11. *Either either either Archibald will marry Adelaide, or Adelaide will elope with Rodney, or the sponsor will cancel the soap opera, or the FCC will force the network to cancel the program.

We see that, in all Chomsky's examples, acceptability is limited to sentences where central self-embedding occurs only once. But the proof requires that there be examples in English where self-embedding occurs to an arbitrary depth. The argument that natural language cannot be produced by a finite system is thus invalid. Being based on assertions that certain sentences are English, while others are not, it fails when the assertions are based on acceptability, because the data do not bear out these assertions. And it fails when the assertions are based on grammaticality, because grammaticality is defined in terms of the system which Chomsky is trying to justify. That is, the proof fails because the reasoning is circular. Chomsky assumes what he sets out to prove.

Postal (1964) attempts a related proof. He claims that Mohawk contains utterances of the form *XX*, where both *X*'s stand for the same string of constituents, and where these strings can be arbitrarily long. Specifically, Mohawk has a structure that allows incorporation of a noun (for example, the direct object of a verb) into the verb. Thus Mohawk has two forms which could be paraphrased *The girl likes the car* and *The girl car-likes*. Postal's argument is based on three assertions. The first is that the acceptable sentences of Mohawk include *The girl car-likes the fast car*; that is, sentences where the incorporated object co-occurs with an identical direct object which has been preceded by a modifier. The second assertion is that there are no Mohawk sentences of this type in which the object of the verb and the incorporated object are not identical; that is, there are no Mohawk sentences of this type which are in the form XY. The third assertion is that the nouns which are used as objects (both incorporated and free-standing) can be arbitrarily long, because of a process of making a noun out of a verb containing an incorporated noun. If this is an accurate description of what happens in Mohawk, these sentences cannot be enumerated by a finite state device, or even in any efficient way by context-free phrase structure rules.[3]

---

[2] A sentence is acceptable to me if my estimate of the probability of occurrence of a sentence of like construction in a natural language text is greater than zero. I exclude from natural language text sentences dreamed up by linguists, psychologists, English teachers and poets.

I have discussed this argument with Floyd Lounsbury. He accepts the first and second assertions with some reservations: there seem to be exceptions to the assertion that the two constructions must be identical. Two examples could be paraphrased *The girl car-likes this* and *The girl corn-feeds the boy*. Another example translates as *People lived there, siblings* – where the word for siblings does not contain the word for *person* morphologically, although one could argue that semantically 'person' is a component of *sibling*. It is not clear how these examples affect Postal's proof. It is the third assertion which Lounsbury finds most difficult to accept. Can incorporated nouns be arbitrarily long? Lounsbury says that you never get more than one incorporation at the creative (syntactic) level. What is incorporated may itself be an incorporation; but if it is, it is a common idiom, a lexeme in the language. The longest such construction of which he is aware in Mohawk is the word for stove polish, which one might paraphrase *house-heater-black-and-shiny-maker*. This example contains only two incorporations. In Oneida the longest such construction he knows about is the verb which means 'to go half way around the altar', where *altar* is a lexeme consisting of the Oneida morphemes for *holy* and *table*, and where *table* consists of the morphemes for *food* and *place*. Thus the maximum of incorporations in Oneida seems to be three. A maximum of two or three incorporations (only one of which is syntactic) seems to be a long way from the arbitrarily large number of incorporations claimed by Postal. If this characterization of Mohawk is accurate, Postal's proof is invalidated.

I assert that not only are these proofs invalid, but that the conjectures are wrong. English (and, I suspect, all languages) can, in fact, be described by a finite state device, namely a network of relationships, where each relationship is itself a finite state device.

Not only has what is possible in natural language been overestimated,

---

3. Bar-Hillel (1967, p. 536) has pointed out that Postal's belief contradicts Chomsky's conjecture (1965, p. 61) that context-sensitive phrase structure rules may be adequate to enumerate the sentences of natural language.

but also the power of a finite state grammar has been seriously underrated. Within relational network theory, the addition of the upward conjunction relation (the 'upward and') allows one to express grammatical information which, within a rewrite rule format, requires not only context-sensitive rules, but also transformations and subcategorization rules (Reich 1970b, 1970c, 1971a, 1971b).

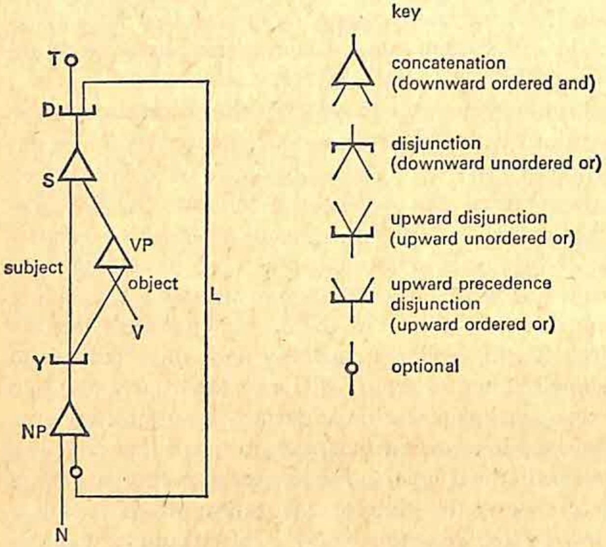

T
D
S
VP
subject
object
L
V
Y
NP
N
key
concatenation (downward ordered and)
disjunction (downward unordered or)
upward disjunction (upward unordered or)
upward precedence disjunction (upward ordered or)
optional

Figure 1

To see how the sentences we have been considering can be described with a finite state grammar, consider Figure 1. This is the network grammar that corresponds to the following rewrite grammar:

$S \rightarrow NP\ VP$
$VP \rightarrow V\ NP$
$NP \rightarrow N\ (S)$

Each of the nodes of the network is a finite state device, and thus can be represented as a finite state diagram. The node representing concatenation is basically a three-state device, as pictured in Figure 2. By definition the node starts out in state 0. When a signal comes into the node on wire a, a signal is sent out wire b and the node changes to state 1. When the structure down path b has been completely realized, a positive feedback signal comes up wire b. When this happens, the node sends a signal down wire c and changes to state 2. When the positive feedback signal comes up c, the node sends a

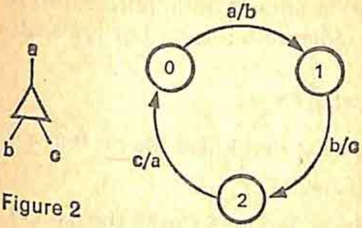


Figure 2

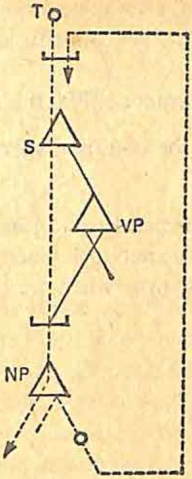


Figure 3

positive feedback signal up wire a and returns to state 0, its initial state. This is what the finite state diagram in Figure 2 says. Three separate states are required to specify the order of the events involved.

So far we have not defined our node so that it can operate properly when there are loops in the system. To see this, consider the path of a signal moving through the network. The heavy line in Figure 3 shows the path of a signal in the case of a sentence with a clause as part of the subject. Note that a signal comes in from the a wire of concatenation element S when it is in state 1. This condition has not been defined. By adding a push-down stack mechanism to our model, we could solve this problem by means of recursion – as does transformational grammar – but we would no longer have a finite system.

There is a better solution. We just add two more states to our definition of concatenation, as shown in Figure 4. This definition is equivalent to adding to each node in the network a push-down stack with a depth limited

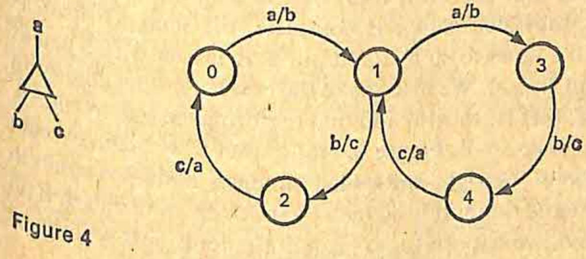


Figure 4

to one. In the case of a sentence modifying the subject noun phrase, this is exactly what we want. We want to allow sentence 5 above, but not sentences like 6 or 7.

But what about sentences like the following?

12. The cow tossed the dog that worried the cat that killed the rat that ate the malt.

Right branching can clearly go on indefinitely. Figure 5 shows the path of the signal through the network when a modifying clause is part of the object. Note that this time, when we arrive at S, it is in state 2 – whereas before, when we had arrived at S from the loop, it had been in state 1. Thus we find that, within the framework of a network of finite elements, there is a quite natural and non-*ad-hoc* differentiation of central embedding from right branching. How do we handle branching to an arbitrary depth? We can ignore the fact that the node hasn't returned to its initial state, since after the right embedded clause is complete, the sentence itself is complete. We return the node to state 1 and put a signal out wire b just as if the node had been in state 0. We also make the corresponding change in the pushed-down case; that is, we also add that performance rule to state 4. This gives us the state diagram in Figure 6. Notice that if we again expand the object noun phrase of the right embedded sentence, node S is again in state 2 and we do the same thing. It is defined by the finite state diagram no matter how many times we choose to go around the loop, just so long as

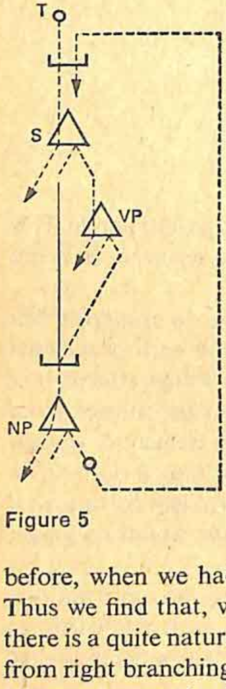


Figure 5

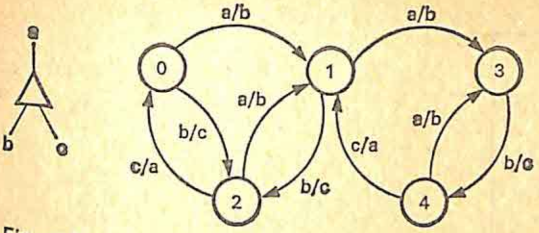


Figure 6

we continue to do it on a right branch. While we are looping, we are not keeping track of the number of times we loop; thus when we send feedback up the path, we cannot send the feedback up around the loop the same number of times – but then we don't really have to. Right branching means that when we are finished with the most deeply embedded construction, we are finished with everything, so we can just send feedback directly to the top of the construction. This is the reason for the precedence disjunction at D in Figure 1. No matter how many times a signal has come into D from loop L, the positive feedback upon completion coming up from S goes directly to T.

So far we see that the relational network model can handle right branching and central embedding, which could also be handled by Yngve's depth hypothesis model (1960). However, this was rightly criticized by Chomsky (1965, pp. 198–9) because it could not handle left branching to an arbitrary depth, which is also acceptable in English. The clearest case of left branching in English is the possessive contruction:[4]

13. Henry is Doug's father's second wife's sister's daughter's husband.

The grammar for this construction would look like Figure 7. A description of the processing will illustrate how left branching can also be handled by the finite state system. The first noun, 'Doug' in our example, is produced as the sole realization of the NQ construction. But upon completion of NQ, instead of going directly up to NP, the positive feedback first goes around loop M, producing the possessive ending and another noun, this time 'father' in our example – and so on around the loop until there are no more possessives to express. At this time a negative feedback signal is sent down from 'Possessive' around loop M to E. When a negative feedback comes in from M to E, then and only then does the positive feedback signal continue to NP.

[4] This is an actual sentence produced by my wife shortly before we were married, when she was learning my family tree. It is a true statement about my family. While this particular sentence may seem unintelligible, the problem lies in the semantics, not the syntax. Try such a sentence on your own family tree.

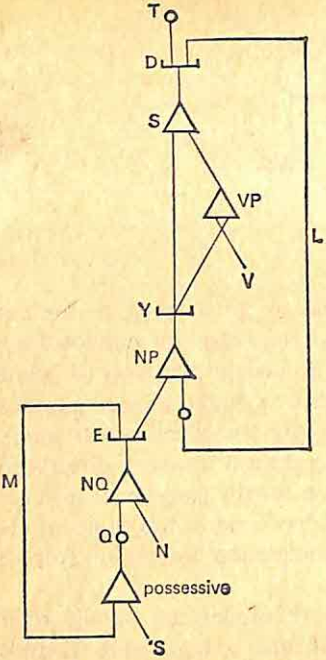


Figure 7

In this type of loop, the concatenation element receives an incoming signal from line b when it is still in state 0. In the same way that we handled right branching, we let the node behave as if the signal had come in when the node was in state 1; that is, we send a signal out c and change to state 2. The resulting definition is shown in Figure 8. This finite addition to the node definition allows left branching an arbitrary number of times. Thus we see that a properly defined finite state relational network can allow left and right branching to an arbitrary depth, while still limiting central embedding so as to agree with what the native speaker finds acceptable.

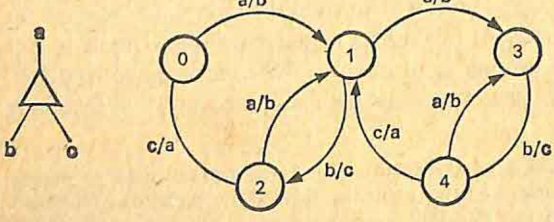


Figure 8

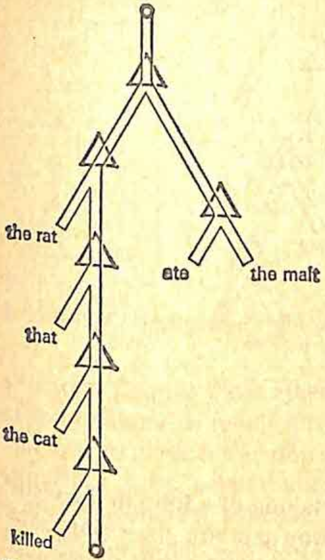


Figure 9

If we carefully study the model just described, we find that it solves several additional problems. Let us look at the derivational structure of a sentence. In this model the derivational history of a sentence is produced by following the signals through the network and copying all the concatenation nodes through which the signals successfully pass. When the signals successfully pass through the same node more than once, that node is copied more than once. Both the regular signal and the positive feedback signal are represented. Using this format, the difference between the recursive generation of sentence structure and the type of generation discussed here is made clear. Figure 9 shows recursive generation: every time a signal goes through a node, the feedback signal must eventually pass through that same node. Figure 10 shows arbitrary right branching: no matter how many times a signal passes through a node, the feedback passes through it only once.

An appropriate term for the type of generation shown in Figure 10 is ITERATION. In computer programming languages, both iteration and recursion have been in use for some time; but because it did not fit in with transformational grammar notation, iterative generation has not been used in linguistics.

Left branching iteration is diagrammed in Figure 11. In this case the signal moves down through the structure only once, but up through the structure an arbitrary number of times.

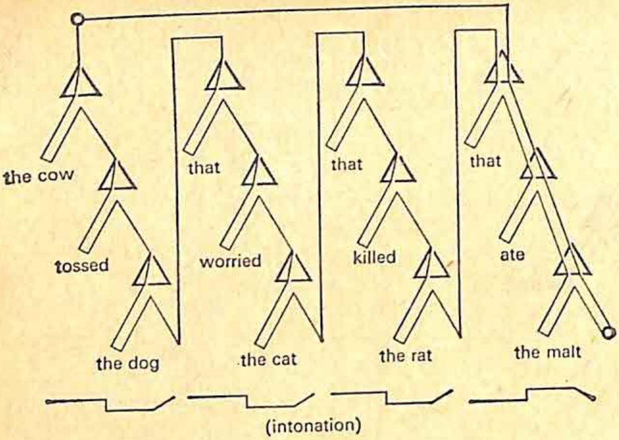


Figure 10 Iteration (right branching)

Note that, in both Figures 10 and 11, the taking of a loop in linguistic structure is marked by a distinctive intonation pattern, the rising non-terminal contour. As shown in Figure 10, this contour begins on the last stressed syllable before the loop, and the rise continues until the loop. When the loop is not taken, the characteristic falling terminal contour results.[5] This intonation structure corresponds quite neatly to looping within syntax, but it does not correspond to transformational surface

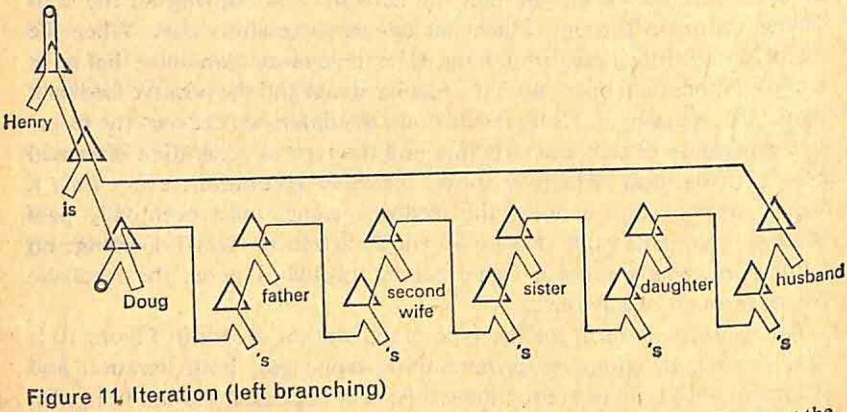


Figure 11 Iteration (left branching)

---

[5] This is, of course, an oversimplification. There is often something special about the last taking of a loop. For example, if there is only one loop, as in *He talked to the girl who lived upstairs*, this intonation pattern normally does not occur. Also, in co-ordination, the *and* or *or* is usually not realized except before the last occurrence of the loop.

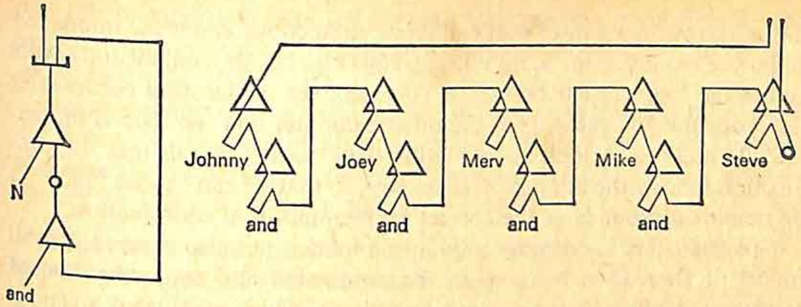


Figure 12 Coordination (right branching)

structure. Chomsky and Halle (1968, pp. 371–2) discuss this problem, and propose either that re-adjustment rules would have to rearrange the surface structure to fit the intonation structure, or else that one should consider the problem a part of performance rather than competence – that is, one should ignore the problem. The relational network solution which I am proposing handles this intonation structure straightforwardly within syntax.

Iteration also solves the problem of how to represent coordination of an arbitrary number of elements with a finite grammatical structure. Within a rewrite rule framework, generation must either be right recursive or left recursive, or else there must be a separate rule-type to handle this type of structure. The problem of right or left recursive structure is that the resulting parenthesization of the sentence has no correspondence in either meaning or phonology. In the relational network iterative approach, the problem is avoided: the derivational history looks like the unstructured list that it is. However, there is the problem that we have two types of iteration, right branching and left branching. Figures 12 and 13 show the two alternative structures. Is this an embarrassment of riches? The answer

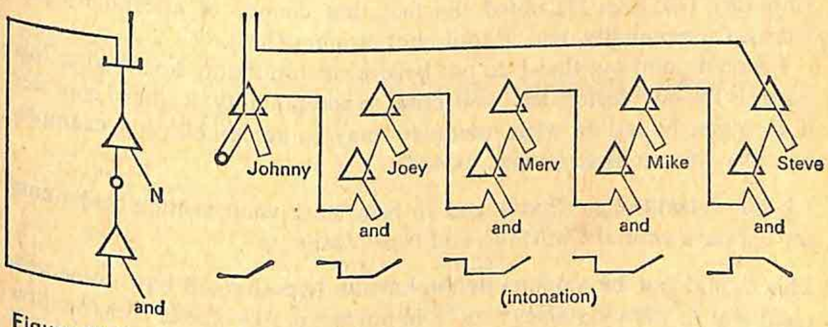


Figure 13 Coordination (left branching)

is no. The two structures make different predictions about the intonation pattern of conjunction. In right branching (Fig. 12), the conjunction occurs before the loop; in left branching (Fig. 13), the conjunction occurs after the loop. But we know from the other examples that the loop is intonationally marked. A look at intonation data quickly reveals that the conjunction follows the intonation boundary, so that we can choose Figure 13 on empirical grounds as the correct representation of coordination.

Note that this is not only a linguistic model, but also a psychological model of short-term memory in the production and comprehension of sentences. Unlike Yngve's depth hypothesis, which postulated a central memory store which could hold about seven items, this model represents memory as storable to a depth of one, all-over linguistic structure. Thus, rather than characterizing memory as narrow and deep, this model characterizes memory as broad and shallow: we might call it a BREADTH HYPOTHESIS of language memory structure.

From this hypothesis one can make quite specific predictions as to what types of constructions should be acceptable and what types should not. For example, the breadth hypothesis would predict that, although nested dependencies of the same item would be unacceptable (as seen in examples 10 and 11), a sentence consisting of nested dependencies of different items should be acceptable. The data seem to bear this out; e.g.

14. If either whenever one nation attacks, the other retaliates, or one nation feels it can destroy the other without suffering serious consequences, then renewal of the war is a logical consequence of those nations' policies.

This sentence seems acceptable in spite of the fact that, from a rewrite-rule standpoint, it has the same dependency structure as sentences 10 and 11, which are not acceptable. The difference, of course, is that the dependencies which must be remembered in sentence 14 are different, and thus the information can be stored in different places in the linguistic structure. Chomsky (1957, p. 23) noted the fact that change of vocabulary can increase acceptability, but he could not account for it.

I should point out that I do not believe the limitation described in this paper is the only factor that contributes to acceptability. Rather I consider it an upper bound on what utterances may be acceptable. For example, consider this sentence (Yngve, p. 457):

15. *He called the girl that he met in New York while visiting his parents for ten days around Christmas and New Year's up.

This would not be stopped by the breadth hypothesis. I have some suspicions as to what variables may be important in this type of construction, but I have not yet run the experiments necessary to test them. I do feel that this phenomenon, like the breadth hypothesis, results from a basic psychological property of language behavior, rather than being some language-specific fact.

One should also bear in mind the fact that people can make estimates of the degree of closeness of an unacceptable string to natural language. McCawley (personal communication) objects to the fact that my proposal implies that the following are equally unacceptable:

16. *That the fact that Max is a homosexual bothers you is obvious.
17. *That that Max is a homosexual bothers you is obvious.

I find these examples equally unacceptable, since I estimate that the probability of occurrence of sentences similar to either one is zero. The fact that 16 is generally preferred over 17 seems to be based on the fact that 18 is preferred over 19:

18. The fact that Max is a homosexual bothers you.
19. That Max is a homosexual bothers you.

That is, each of the unacceptable sentences contains an acceptable portion, and one of these portions is stylistically preferred over the other.

In this paper I have attempted, first, to show that Chomsky's and Postal's proofs that language cannot be produced by a finite system are invalid. Second, I have suggested that the conjecture is wrong, and I have proposed a finite system which I believe to be adequate. This system generates by means of recursion to a depth of one, and by means of iteration an arbitrary number of times. I have tried to show that this model accounts for data which cannot be accounted for within a rewrite-rule framework – namely, central embedding versus right and left branching; the intonation pattern of sentences containing strings of relative clauses; multiple branching constructions such as coordination; and acceptability of nested dependencies, provided the dependencies come from different constructions.

References

BAR-HILLEL, Y. (1967), 'Review of the structure of language', in J. Fodor and J. Katz (eds.), *Language*, vol. 43, pp. 526–50.

BEVER, T. G., FODOR, J. A., and GARRETT, M. (1968), *A Formal Limitation of Associationism*, in T. R. Dixon and D. L. Horton (eds.), *Verbal Behavior and General Behavior Theory*, Prentice-Hall, pp. 582–5.

CHOMSKY, N. (1956), 'Three models for the description of language', *IRE Trans. Information Theory*, IT-2, pp. 113–24.

CHOMSKY, N. (1957), *Syntactic Structures*, Mouton.

CHOMSKY, N. (1965), *Aspects of the Theory of Syntax*, MIT Press.

CHOMSKY, N., and HALLE, N. (1968), *The Sound Pattern of English*, Harper & Row.

MARKS, L. E. (1968), 'Scaling of grammaticalness of self-embedded English sentences', *Verbal Learning and Verbal Behavior*, vol. 7, pp. 965–7.

MILLER, G. A. (1962), 'Some psychological studies of grammar', *Amer. Psychol.*, vol. 17, pp. 748–62.

POSTAL, P. M. (1964), 'Limitations of phrase structure grammars', in J. Fodor and J. Katz (eds.), *The Structure of Language*, Prentice-Hall, pp. 137–51.

REICH, P. A. (1970a), *A Relational Network Model of Language Behavior*, University of Michigan Ph.D. dissertation, University Microfilms: Ann Arbor.

REICH, P. A. (1970b), 'Relational networks', *Canad. J. Linguistics*, vol. 15, pp. 95–110.

REICH, P. A. (1970c), 'The English auxiliaries: a relational network description', *Canad. J. Linguistics*, vol. 16, pp. 18–50.

REICH, P. A. (1971a), 'Competence, performance, and relational networks', in A. Makkai and D. Lockwood (eds.), *Readings in Stratificational Linguistics*, University of Alabama Press.

REICH, P. A. (1971b), 'Symbols, relations and structural complexity', in A. Makkai and D. Lockwood (eds.), *Readings in Stratificational Linguistics*, University of Alabama Press.

YNGVE, V. H. (1960), 'A model and an hypothesis for language structure', *Proc. Amer. philos. Soc.*, vol. 104, pp. 444–66.

Part Four
Experiment and Discovery

Techniques for revealing the hidden structure of language were especially loved by American linguists of the 1950s, but are in some measure a mark of all the many schools referred to as 'structuralist'. Each of these Readings describes or illustrates a methodology – as do, in part, the earlier papers of Hatcher, Waterhouse, Yngve, and Pike, to name only a few (Readings 3, 4, 8, and 14). All these methods have the added charm of *in toto* rejection by Chomsky, Postal, and other transformationalists – statistics, co-occurrence frame tests, informant response tables (though these are sometimes used by transformational grammarians).

Vietnamese is the only 'isolating' language discussed in this collection, Eskimo and Aleut are among the most polysynthetic of all languages, and the deaf-mute manual language studied by Tervoort is sometimes denied the status of natural language. It is especially interesting to compare its syntax with those appearing in the experiment involving the chimpanzee Washoe (see Gardner and Gardner, 1969). The experiment with the chimpanzee Sarah (see Premack, 1970) is not quite as revealing. The suggestion is inescapable that some linearities of spoken natural languages are direct consequences of the use of sound as a carrier, and need not be wired into the human brain.

References

GARDNER, R. A., and GARDNER, B. T. (1969), 'Teaching sign language to a chimpanzee', *Science*, vol. 165, pp. 664–72.
PREMACK, D. (1970), 'The education of S*A*R*A*H', *Psychology Today*, vol. 4, pp. 54–8.

20 P. J. Honey

Word Classes in Vietnamese

P. J. Honey, 'Word classes in Vietnamese', *BSOAS*, vol. 18, 1956, pp. 534–44.

The phonological structure of the Vietnamese word as used in the system of Vietnamese word classes established in this paper may be stated in terms of syllables. The syllable is defined as the element of Vietnamese having the prosody of one, but not more than one, of the system of six tones set up for the analysis of the Vietnamese language at the phonological level. The syllabic unit is used to refer to an element of text at the grammatical level of analysis corresponding to the syllable at the phonological level. A study of Vietnamese texts enables the following statements to be made:

1. The great majority of the syllabic units in Vietnamese may be described as positionally free. By positionally free is meant that the syllabic units referred to are not restricted to positions immediately preceding or following one or more of a very limited number of other syllabic units. Syllabic units which are so restricted are referred to as bound.

2. A small number of syllabic units are bound, being found only in conjunction with one or more of a very restricted number of other syllabic units. Most of such bound units are found in conjunction with only one other positionally free unit, but in rare cases are found in conjunction with two or more.

An example of a bound unit is nhách, which is found only when directly preceded by the syllabic unit dai. This disyllabic unit dai-nhách is positionally free, as is the syllabic unit dai.

e.g. Thịt gì mà lại *dai-nhách* thẽ này?
What is this meat that it tastes so very tough?
Dai như chão rách
Very stubborn

In order to set up the system of word classes described in this paper it is necesssary to redefine the unit 'word' in Vietnamese and not to equate it with the 'single morpheme' or the 'monosyllable' as has been attempted in the past. Such equations have not yielded satisfactory results.[1] The Vietnamese word is, therefore, redefined as the positionally free syllabic unit or the minimal sequence of such units of which at least one unit is bound. Where it is necessary to distinguish between the monosyllabic and the polysyllabic word, the term 'simple word' is used to describe the former and 'complex word' to describe the latter.

In the assigning of words to their word classes, complex words are treated as words entirely separate from, and having no relationship with, the syllabic units of which they are constituted, even when these syllabic units may themselves appear to be referable to simple words. The following analogy taken from English will serve to illustrate this point. When assigning English words to grammatical categories, the grammarian will treat the two words 'loganberry' and 'berry' as being separate from and independent of each other. The relationships which could be stated at other levels of discourse are irrelevant at this level of abstraction. Bound units of Vietnamese cannot be classified since they are not treated as words in this analysis.[2] Similarly, the bound form 'logan' cannot be assigned to a word class in English since it is found only as an integral part of a complex word and never in isolation.

It would have been possible to divide the words of Vietnamese into a greater or lesser number of categories. The actual number of word classes will depend upon the nature and the number of the criteria employed in the establishment of the classes. The object of the grammarian is to establish categories of word which are most efficient for enabling him to make the most comprehensive and, at the same time, the simplest grammatical statements about Vietnamese. It is not claimed that the Vietnamese word classes established in this paper are the 'true' or the only classes of word which may be set up for the Vietnamese language, and they may be superseded whenever a more efficient system of classes is established[3]. They do, however, enable systematic syntactic statements to be made and are themselves wholly definable in terms of such statements.

---

[1] In his book (1951, p. 44, 2.1) Professor M. B. Emeneau emended the definition of the Vietnamese word slightly when he stated: 'The basic unit of the syntactic analysis of the language is the word which is the phonological unit and, at the same time, the morphological unit as was stated in 1.1. The word and the morpheme coincide, with the qualification made in 2.4.' The qualification referred to was concerned with reduplication in groups of two morphemes where there was a partial similarity of form between the two.

The Vietnamese scholar Dr Lê văn Lý (1948, p. 133), redefined the Vietnamese word and established two types of word, '*mot simple*' and '*mot composé*'. The *mot simple* is defined as 'Un signe vocal dont la formation peut partir d'un simple phonème ou de la combinaison de plusieurs phonèmes, dont l'expression vocale se fait en une seule émission de voix ou une seule syllabe que l'écriture présente par une unité distincte et qui offre un sens intelligible'.

The *mot composé* is defined as a unit of more than one syllable, substitutable for a *mot simple*, and expressing a single concept.

It is clear from this that neither Emeneau nor Lê văn Lý was satisfied with the usual definition of the Vietnamese word.

[2] Emeneau did attempt such a classification (1951, p. 79, 2.14) because, according to his definition, such units are words, but the following quotation shows the difficulty into which such an attempt leads. 'An example is hấu, which occurs only as an attribute following the noun dưa which denotes any species of cucurbitaceous plant, including various types of melon and cucumber. The phrase "dưa-hấu" means "watermelon". It is impossible to assign hấu to the verb or to the noun class, and the decision can be suspended. The parallels dưa-chuột "cucumber" in which chuột is the noun "rat", and dưa-đỏ, also "watermelon", in which đỏ is the verb "to be red", hardly help the analysis.'

The word classes here established are twelve in number. The words which make up each class have certain features in common which are not shared by the words of other classes. The classes themselves are numbered from 1 to 12.

Word class 1

Word class 1 is described by means of the positional relationships of all its members with four selected Vietnamese words. These four words do not themselves form a word class and are, indeed, members of different word classes. They are simply four Vietnamese words which are selected as criteria for the setting up of the first class in this system of word classes.

These four words are nhiều, lắm, rất and hơi, and they are used in the following manner. All Vietnamese words which are found directly to precede, but never to follow either or both of the words nhiều and lắm are selected to form a set of words comprising word class 1 and word class 2.

e.g. Tôi đọc nhiều sách.
I, read, many, book.
I read a great many books.

Since the word đọc is never found to follow the word nhiều or the word lắm directly but is, as may be seen from the above example, found to precede nhiều, đọc is selected as a member of this set of words.

e.g. Ngôi nhà này hẹp lắm.
wing, house, this, small, very.
This house is very small.

---

[3] Some of the criteria used in the classification of words described in this paper resemble those set out by Dr Lê văn Lý in his chapter dealing with word classification (1948, p. 145). The Vietnamese word as defined in this paper is very different from the word defined by Dr Lê văn Lý. The number of word classes established in this paper and the content of these classes are different from those of Dr Lê văn Lý.

The word hẹp in the above example is never found to follow directly either the word nhiều or the word lắm and is therefore selected as a member of the set of words described above.

The isolation of class 1 words from this set is effected by the use of the words rất and hơi. All words of this set which are found to follow directly either rất or hơi are assigned to word class 1.

e.g. Ngôi nhà này rất hẹp.
wing, house, this, very, small.
This house is very small.

In the earlier example the word hẹp was shown to be a member of the set of words comprising word classes 1 and 2. In this example the word hẹp is found directly to follow rất. In the light of this evidence the word hẹp is assigned to word class 1.

e.g. Quyển sách này hơi đắt.
volume, book, this, slightly, expensive.
This book is a little expensive.

Quyển sách này đắt lắm.
volume, book, this, expensive, very.
This book is very expensive.

The word đắt in these examples is assigned to word class 1.

Word class 1 is thus established by means of these four selected Vietnamese words.

Word class 2

The second Vietnamese word class is described by means of the same four selected words used in the establishment of word class 1. It has already been shown how a set of words was established, the members of which were either class 1 or class 2 words. The establishment of this set of words entailed the use of the two selected words nhiều and lắm. Subsequently class 1 words were separated from the set by means of the two selected words hơi and rất. The remaining members of this set of words, after the separation of class 1 words, are class 2 words.

All Vietnamese words, therefore, which are found directly to precede but never to follow either or both the words nhiều and lắm, and which are never found directly to follow either hơi or rất, are assigned to word class 2.

e.g. Tôi đọc rất nhiều sách.
I, read, very, many, book.
I read a very large number of books.

The word đọc, which has already been established in an earlier example as a member of the set of words comprising class 1 and class 2 words, is never found directly to follow either rất or hơi. It is therefore assigned to word class 2. It is of interest to note that the word nhiều itself directly follows rất in the foregoing example and is assigned to word class 1.

e.g. Anh ấy đã xem nhiều người lính đi qua đây.
brother, that, past, see, many, man, soldier, go, pass, here.
He saw a number of soldiers going past.

The word xem, which is never found directly to follow either rất or hơi, is seen in the foregoing example to precede nhiều. Xem is therefore assigned to word class 2.

Two word classes are thus established by means of four selected Vietnamese words. Though these four words have places in the structural order of the sentence as members of word classes, they are here used on their face value as criteria which have been proved useful for the establishment of the twelve word classes. It was decided to designate words so used as 'indicators'.[4] The term indicator is used throughout the remainder of this paper to refer to any Vietnamese word selected as a criterion with the aid of which a word class is established.

Word classes 3, 4 and 5

Indicators are also used in the establishment of word classes 3, 4 and 5. All Vietnamese words which are found:

(a) not to be members of word classes 1 or 2, and

(b) to be directly preceded by any one of the following indicators, các, những, mấy, nhiều, bao-nhiêu, or

(c) to be directly followed by one of the indicators này, ấy, kia,

are assigned to a set of words, the members of which comprise word classes 3, 4 and 5.

In the following examples each sentence contains one or more of the indicators used in the establishment of this set of words. The words which are not members of either word class 1 or 2, and which stand in the prescribed positional relationship to these indicators are printed in italics.

e.g. Cái *bàn* này giá bao-nhiêu tiền?
object, table, this, cost, how much, money.
How much does this table cost?

---

[4] I have adopted the term 'indicator' from the language of chemistry in which it describes a reagent used to indicate by change of colour, the presence of an acid, alkali, etc.

Mấy *tháng* nữa tôi sẽ lên Hà-nội.
several, month, more, I, future, go up, Hanoi.
I shall go to Hanoi in a few months' time.

Về mùa xuân cái *vườn* kia đẹp quá.
return, season, spring, object, garden, that, beautiful, too much.
That garden is very beautiful in the spring.

The division of this set of words into its three constituent word classes is effected by means of the indicator một. Certain members of this set of words are found to be directly preceded by the word một, while others are never found in this position. The members of the set of words under investigation are printed in italics in the following examples.

e.g.	Một *ngày*	Một *tháng*
	one, day	one, month
	One day	One month

Tôi ở đây một *tháng* rồi.
I, at, here, one, month, already.
I have been here for a month.

Nó làm xong công-việc ấy trong một *ngày*.
he, do, finish, work, that, inside, one, day.
He finished the job in one day.

Trong buồng kia chỉ có một cái *ghế* thôi.
inside, room, that, only, have, one, object, chair, stop.
There is only one chair in that room.

Cha tôi có một cái *nhà*, bốn mẫu ruộng, và một con *trâu*.
father, I, have, one, object, house, four, acre, ricefield, and, one, creature, buffalo.
My father owns a house, four acres of ricefield, and a buffalo.

In the two latter examples it will be observed that the words italicized are not directly preceded by the indicator một in a two-word group, the Vietnamese for 'one house', 'one chair', and 'one buffalo' being the three-word groups một cái nhà, một cái ghế, and một con trâu.

The members of the set of words are divided into word classes 3, 4 and 5 in the following way.

Word class 3

All members of the set of words being investigated which are never found to be directly preceded by the word một are assigned to word class 3.

Word class 4

All members of the set of words being investigated which are found in the intermediate position between the word một and the following class 3 word, are assigned to word class 4.

Word class 5

All members of the set of words being investigated which are found to be directly preceded by the word một and which are never followed by a class 3 word are assigned to word class 5.

It will be seen that, of the members of the set of words being investigated which are cited in the foregoing examples, nhà, ghế, and trâu are assigned to word class 3, cái and con (which are not printed in italics) to word class 4, and ngày and tháng to word class 5.

Word classes 6 and 7

Indicators are not used in the establishment of these word classes. The classes are set up by means of colligation with word classes 3, 4 and 5.[5] All Vietnamese words which are found directly to precede words of class 5, class 4, or groups formed by a class 4 word followed by a class 3 word when these groups are substitutable for single class 5 words, are assigned to a set of words comprising word classes 6 and 7. In the following examples members of this set of words are printed in italics.

(i) Followed by class 5 word.

Ba tháng
three, month
Three months

Tháng sau anh ấy về nhà.
month, after, brother, that, return, house.
The following month he returned home.

Ba tháng sau anh ấy về nhà.
three, month, after, brother, that, return, house.
Three months later he returned home.

(ii) Followed by class 4 word.

Sáu con
six, living creature
Six animals

---

[5] For 'colligation' see H. F. Simon (1953).

Anh ấy có mấy con trâu? Tôi có *sáu* con.
brother, that, have, how many, living creature, buffalo, I, have, six, living creature.
How many buffaloes has he got? I've got six.

Class 4 words are never found unless they are either preceded by a member of the set of words being investigated or else followed by a class 3 word. It is therefore not possible to provide an example of a class 4 word without a preceding member of the set of words under investigation and without a following class 3 word.

(iii) Followed by the group 'class 4 followed by class 3'.
Những cái bàn
plural, object, table
The tables

Cái bàn này làm bằng gỗ.
object, table, this, make, by means of, wood.
This table is made of wood.

Những cái bàn này làm bằng gỗ.
plural, object, table, this, make, by means of, wood.
These tables are made of wood.

The division of this set of words being investigated into its two constituent word classes, class 6 and class 7, is effected by means of the indicators độ, chừng, and thứ in the following way.

Word class 6

All members of the set of words under investigation which are found to be directly preceded by any one of the indicators độ, chừng, and thứ are assigned to word class 6.

Word class 7

All members of the set of words under investigation which are never found to be directly preceded by any of the indicators độ, chừng, or thứ are assigned to word class 7.

e.g. Độ *chín* người đã đến.
approximately, nine, person, past, come.
About nine people came.

Lần này là lần thứ *năm*.
time, this, – , time, order, five.
This is the fifth time.

In the two preceding examples the words chín and năm are assigned to word class 6 because they are directly preceded by the two indicators độ and thứ respectively. The word những, however, illustrated in an earlier example, is never directly preceded by any of the indicators and is assigned to word class 7.

It may be of interest to note that the indicator một which was used in the setting up of word classes 3, 4 and 5, is itself a member of word class 6.

Word class 8

The words which make up the eighth word class are very few in number – there are only 11 in all. They enter into definite syntactic relationships by virtue of which they are grouped together. For example, the words of this class enter into many of the syntactic relationships of words of classes 3, 4 and 5, yet, unlike words of these classes, they are never preceded by words of classes 6 and 7. They cannot be classified as class 3, 4 or 5 words by means of the indicators used in setting up those categories. The word chúng and all other Vietnamese words which, when directly preceded by the word chúng, form endocentric groups substitutable in all contexts for the single words of the class, are assigned to word class 8. The following is a complete list of the members of word class 8.

Tôi : ta : tao : mình : mày : bay
I we I I, myself you (singular) you (singular)
người : nó : hắn : tớ : chúng
you he, she, it he, she I they, plural marker.

e.g. *Tôi* đi xem xi-nê mới về.
I, go, see, cinema, recently, return.
I have just come back from the cinema.

Chúng không đồng ý với *ta* về vấn-đề này.
they, not, same, idea, with, we, about, question, this.
They disagree with us on this question.

Chúng ta sẽ gặp nhau khoảng ba giờ chiều.
plural we, future, meet, one another, approximately, three, hour, afternoon.
We shall meet at about 3 p.m.

It would be possible to substitute either words of classes 3, 4 and 5 or endocentric groups consisting of a word of class 6 or 7 followed by a word of class 3, 4 or 5 for any of the class 8 words in the above examples. Words of classes 6 and 7 would never be found, however, to precede directly any of the class 8 words.

Word classes 9, 10, 11 and 12

All Vietnamese words which are not members of word classes 1 to 8 are assigned to a set of words which comprises word classes 9, 10, 11 and 12. Members of this set of words are assigned to their respective word classes according to the place or places in which these words are found in the clause. Clause, in this context, is defined as being the smallest connected Vietnamese utterance bounded by pause or potential pause. Some members of this set of words are found only in the initial place of the clause, some only in a medial place, some only in the final place, while some are found in more than one of these places. The following word classes are established.

Word class 9	Words which are not members of word classes 1–8 and which are found only in the initial place of any clause.
Word class 10	Words which are not members of word classes 1–8 and which are found only in a medial place of any clause.
Word class 11	Words which are not members of word classes 1–8 and which are found only in the final place of any clause.
Word class 12	Words which are not members of word classes 1–8 and which are found in more than one of the three places in a clause described above.

Typical words of these four classes are:

Word class 9	nếu, hễ, dẫu, hằng
Word class 10	hoặc, hay, với
Word class 11	ạ, ơi, nhé, nhỉ
Word class 12	ngay, rồi, luôn, mà

The following are some examples of words of these four classes in Vietnamese utterances. In each example the word of class 9, 10, 11 or 12 is printed in italics.

Word class 9 (initial)

Nếu tôi đi chợ thì anh phải ở lại trông nhà.
if, I, go, market, then, brother, must, stay, watch, house.
If I go shopping then you must stay and look after the house.

Word class 10 (medial)

Mời ông xơi cơm *với* tôi.
invite, grandfather, consume, rice, with, I.
Won't you come and have dinner with me?

Word class 11 (final)

Ta đi xem xi-nê *nhé*?
we, go, see, cinema, – .
Shall we go to the pictures?

Word class 12 (polytopic)

> Ánh đã ăn chưa? Ăn cơm *rồi.*
> brother, past, eat, not yet, eat, rice, already.
> Have you had dinner yet? Yes, I've dined already.
>
> Trước hết tôi đi xem xi-nê, *rồi* về nhà.
> before, complete, I go, see, cinema, already, return, house.
> First of all I went to the pictures, and then I went home.[6]

In the past, considerable difficulty has been experienced by scholars in the investigation of Vietnamese at the grammatical level. This appears to have been due to the fact that no satisfactory statement of word classes had ever been made. It is claimed for the system of word classes described in this paper that, since the criteria employed are formal, even mechanical, it should be possible for scholars to assign the same Vietnamese words to the same word classes. This has not been the case when they have relied upon more notional criteria for establishing categories of word. In addition, the categories established in this paper may be stated in unambiguous terms.[7]

---

[6] It is necessary to point out that difficulties may appear to arise in the case of homophones. For example, in the following utterances:

Để cái cưa này trên cái bàn.
place, object, saw, this, upon, object, table.
Put this saw on the table.

Tôi phải cưa nhiều cây trong khu này
I, must, saw, many, tree, within, region, this.
I have to saw down a large number of trees in this region.

the unit cưa must be assigned to two different word classes. In the first utterance cưa directly precedes the indicator này and is assigned to word class 3. In the second utterance cưa directly precedes the indicator nhiều and must be assigned to word class 2. It is stated in all such cases that the unit in question represents more than one Vietnamese word, and that these words are homophones. In the examples given, the unit cưa represents two Vietnamese words. Indeed, the English translation of cưa, namely 'saw', may be a member of both nominal and verbal categories: to saw: the saw.

A further parallel from English may serve to exemplify this apparent difficulty. It is possible to distinguish homophones in English by means of scatter. The English noun 'ride' has a scatter of 'ride, rides', and may be distinguished from the English verb 'ride' which has a scatter of 'ride, rides, rode, riding, ridden'. Vietnamese is not an inflecting language and it is not possible to distinguish Vietnamese homophones by this criterion of scatter.

[7] Dr Lê văn Lý pointed out the advantages of such a method of word classification (1948, p. 146, 1.2), but himself resorted to subjective criteria in the establishment of his categories as the following extract shows: 'Sont identifiés comme A tous les mots

Since it is difficult to remember numbers, it may be found convenient to employ conventional grammatical terminology to describe the word classes established in this paper. The following table sets out a list of terms which I have found useful in my study of Vietnamese grammar.

Word class 1	Adjective
Word class 2	Verb
Word class 3	Qualified noun
Word class 4	Qualifier
Word class 5	Unqualified noun
Word class 6	Numeral
Word class 7	Marker of plurality
Word class 8	Personal pronoun
Word class 9	Initial particle
Word class 10	Medial particle
Word class 11	Final particle
Word class 12	Polytopic particle

These suggested terms are no more than mnemonics for the word classes established in the paper. None of them should be taken as carrying the connotations which similar terms have in traditional grammar.

---

précédés d'un classificateur tel que cái, con, người, ou d'un mot générique comme kẻ, sư, đồ, việc, nghề.'

The categories *classificateur* and *mot générique* have nowhere been established by Dr Lê văn Lý and it is very doubtful whether any two scholars would agree on the question of which Vietnamese words are *classificateurs* and which *mots génériques.*

References

EMENEAU, M. B. (1951), *Studies in Vietnamese (Annamese) Grammar*, University of California Press.

SIMON, H. F. (1953), 'Two substantival complexes in standard Chinese', *BSOAS*, vol. 15, no. 2, pp. 327–55.

VĂN LÝ, LÊ (undated; preface 1948) *Le Parler Vietnamien*, Éditions Huong Anh, Paris.

21 P. L. Garvin

A Study of Inductive Method in Syntax

P. L. Garvin, 'A study of inductive method in syntax', *Word*, vol. 18, 1962, pp. 107–20. Reprinted in *On Linguistic Method*, Humanities Press 1964, pp. 63–77.

This paper constitutes an attempt to render explicit the procedure followed intuitively in the conduct of the syntactic analysis of a little-known language. It is assumed that the detailed explication of the analyst's intuitive research decisions will ultimately give a precise and reliable understanding of the process of analysis, thus leading to the formulation of an inductive research methodology for syntax.

The emphasis in this approach is thus on the development of a method rather than of a form of statement. The contention is that the application of a controlled procedure will yield reliable results. The form of the statement will then depend on the nature of the results.

I

Linguistic analysis is essentially an inductive process, in the sense that it attempts to derive a listing of elements and a set of statements from either an examination of the responses of informants or the study of texts. It is based on the assumption that in both of these sources of data it is possible to discern regularly recurrent elements of different types and orders of complexity. The classification of these elements and the statement of their conditions of distribution (= patterns of co-occurrence) resulting from the analysis is then considered to constitute an inductive description of the language.

The techniques of linguistic analysis, broadly speaking, fall into two categories: those drawing upon the co-variance of form and meaning in linguistic structure, and those drawing upon the regularities of the occurrence patterns of linguistic elements. The former can be called form-meaning techniques; the latter are usually called distributional techniques.

The total process of linguistic analysis cannot yet be characterized in sufficient detail to produce a complete set of exact analytic rules that are universally and mechanically applicable to any language whatsoever. It is, however, possible to stipulate some of the fundamental features of this process and to formulate gross rules which can be adapted to the conditions of a particular language. It is also possible to foresee that an application of these gross rules to a variety of languages will lead to their further refinement and that a degree of universality is achievable on the basis of inductive generalizations from this varied application.

The basic features of the analytic process can be deduced from a set of assumptions regarding the function of natural language (it is a system of signs) and its multidimensional character (the various 'levels').[1] The process as a whole can be grossly characterized as a progression by series of iterative cycles, each series roughly corresponding to the various presumed levels of the language.

A first approximation of the sequence of procedural steps composing each analytic cycle can be outlined as follows:

1. Formulation of the immediate analytic objective;
2. Preparation of a data base by elicitation, study of text, or reorganization of existing data;
3. Impressionistic examination of data to observe pertinent units and relations;
4. Operational tests when necessary to verify impressionistic observations, leading to attestation of relations and definition of units;
5. Collection and examination of additional data to cross-check relations and definitions.

In any given cycle, a particular step may become inapplicable, either because the next preceding step has failed to yield satisfactory results and a new cycle must be initiated with a reformulated objective, or because the results of the next preceding step or cycle are so satisfactory that the given step has become unnecessary.

There is a significant difference between the analytic cycles of morphemics and phonemics. In phonemic analysis, the boundaries of fused units are directly observable on the basis of breaks in the continuity of the stream of speech. These breaks can be formalized as junctures and then serve as defining criteria for fused units.[2] In morphemics, no comparable direct evidence of fused-unit boundaries exists; they have to be inferred by the application of suitable form-meaning and distributional techniques.

---

[1] Language is assumed to have three sets of levels: (1) two levels of structuring, the phonemic and the morphemic; (2) two levels of organization, selection and arrangement; (3) several levels of integration for both the phonemic and the morphemic level of structuring to account for the progressive fusion of elementary units (phonemes and morphemes) into fused units. For a detailed discussion of this concept of levels, see Garvin (1963), pp. 3–22.

[2] For a detailed discussion of phonemic procedure, see Paul L. Garvin and Madeleine Mathiot (1958).

II

In an earlier paper I have discussed the application of such a sequence of procedures to morphological analysis (Garvin, 1957). In the present paper, I should like to discuss the possibility of establishing a comparable sequence of operations for syntactic analysis. This discussion is based on my experience with the morphemic analysis of Ponapean, a Micronesian language of the Eastern Carolines.

Morphology is here divided from syntax in the conventional way: morphology is the study of the internal structure of words. Syntax is the study of the external functioning of words. The morphology-syntax division thus presupposes that words be properly defined as linguistic units.[3]

In the succession of analytic cycles in morphology, form-meaning techniques are used to ascertain the boundaries and variants of elementary units (morphemes). Distributional techniques are used to classify these units and to ascertain the patterns of integration into units of a higher order of complexity such as morpheme clusters and words. The form-meaning techniques of morphological analysis are based upon the elicitation of morphological paradigms which consist essentially of ordered sets of informant words obtained by a controlled questioning procedure. Distributional techniques in morphological analysis are based upon the observation and experimental manipulation of the occurrence pattern of morphemes within the informant words.

The first purpose of distributional analysis in morphology is to arrive at a properly defined distribution frame, namely, the word (properly defined, as opposed to the informant words which are the behavioral units resulting from elicitation). Within the word, it is possible to define classes of morphemes and morpheme clusters. In languages presenting consistently ordered strings of morphemes and morpheme clusters, it is possible to define the word by using this fixed order as a criterion, and to test for it by a variant of the permutation test, the inversion test. This consists in ascertaining whether or not any pair of adjacent morphemes or morpheme clusters within an informant word can be inverted: if the inversion test fails, the pair belongs to the same word; if the inversion test is positive, each morpheme or morpheme cluster of the pair constitutes a separate word.

The input to the morphological analysis is thus a set of behavioral units, namely, informant responses elicited by a form-meaning technique in ordered sets under controlled conditions. These 'informant words' constitute the test frames for the application of distributional techniques. The output of the morphological analysis is a set of linguistic units, namely, morphemes and their allomorphs, morpheme clusters, and words, as well as the classification or partial classification of these units. This output, in turn, serves as input to the subsequent analysis, that of syntax.

---

[3] '. . . a precise statement of morpheme distribution requires as its proper distributional framework a larger morphemic unit, within the boundaries of which the occurrence of individual morphemes can be rigorously described. . . . If the definition of such a frame unit can be based on criteria of a different sort than the definition of any other larger morphemic unit, such a morphemic unit can properly be called a word. Conversely, if no such definition of a larger morphemic unit is possible, it can be said that the language in question has no words in our sense' (Garvin, 1957, p. 21).

III

I have elsewhere discussed some formal relations and outlined some distributional techniques that can serve in syntactic analysis (Garvin, 1958). The relations are those of agreement, government, and various dependences. The techniques are those serving to ascertain the above relations: in particular, morphological inspection to ascertain agreement and government, the 'dropping test' (see footnote 4) to ascertain dependences, as well as the well-known substitution technique.

I should now like to discuss the role of these relations and techniques in the process of syntactic analysis as applied to Ponapean, as well as point out what additional relations and techniques are required to complete the analysis.

In Ponapean, the word is defined by the fixed-order criterion cited above. The morphological analysis has yielded the following output relevant to the subsequent analysis of the syntax: word boundaries, and a partial classification of words based on their internal structure. This classification consists in the subdivision of words into two major word types, inflected and non-inflected (the latter are further subdivided into quasi-inflected, pseudo-inflected, and uninflected). Inflected, quasi-inflected, and pseudo-inflected words are further classified into morphologically defined word classes (the inflected word classes are nouns and verbs); uninflected words can at this stage not be further classified. The ultimate aim of the syntactic analysis is to define syntactic units, ascertain their relations, and, as a by-product, complete the classification of words by assigning the residue of uninflected words to appropriate syntactically defined word classes.

The cross-checking step of the final morphological cycle has resulted in the marking of word boundaries and the indication of morphological word-class memberships in the data at hand. The data themselves at this point largely consist of elicited informant words and examples culled from text on the basis of their impressionistic similarity to informant words.

The immediate objective of the first syntactic cycle can be stated in terms of the underlying fused-unit conception of language: it is to determine the lowest order of syntactic fused units on the basis of some pertinent defining criterion. The morphologically defined word classes resulting from the previous analysis will allow a more detailed working assumption: that inflected words, namely, verbs and nouns, may well be basic constituents of such lowest order syntactic units as phrases.

The data base of the first syntactic cycle is obtained by separating in the data at hand one-word examples from all those examples which contain nouns and/or verbs and consist of more than one word each. The latter can be searched for the presence of likely syntactic relations, as well as be compared to one-word examples. In the examination step, this search is conducted in the order of tractability established by previous research: agreement, then government, then dependences.[4]

An inspection of the inflectional morphemes contained in the selected examples will bear out the impression created by the inflectional morpheme classes as a whole; namely, that the relations of agreement and government are in Ponapean limited to a few special cases. The data must therefore be examined for the likelihood of the set of relations next in order of tractability, the dependences, beginning with the most tractable of these – unilateral occurrence dependence.

---

[4] 'Agreement and government, once known, can be established by a simple inspection of the paradigmatic morphemes of a given textual segment without the need for further evidence or testing. If not yet known, the existence of agreement can be ascertained from as little as a single pertinent example. The existence of government, since the paradigmatic sets of the governing and the governed do not necessarily share a grammatical category by which the relation can be easily recognized, requires several examples with the governing word held constant and the governed words varying; government can then be said to exist if these various governed words all contain at least one paradigmatic morpheme each of the same grammatical category.

Thus, if agreement and government can be found to constitute the defining criteria of syntactic units, the definition and delimitation of these units can be carried out by a relatively simple inspection procedure. More elaborate analytic procedures are necessary for languages in which agreement and government are trivial or absent. . . .

Dependences can be defined as various relations of presupposition. I have found three types of dependence relations analytically useful so far: unilateral dependence, bilateral dependence and mutual tolerance.

Unilateral dependence is a relation in which A presupposes B but not vice versa.

Bilateral dependence is a relation in which A presupposes both B and C but neither B nor C presupposes A.

Both of the above fall under the relation called "determination" by Hjelmslev.

Mutual tolerance is a negative dependence: it is the relation where, while A and B are allowed to occur together, neither presupposes the other. This is the relation called "constellation" by Hjelmslev.

I found, however, that the relation of unilateral dependence had to be further specified. On one syntactic level of integration, it could be said that unit A presupposes unit B *for its occurrence*; this type of dependence I have called occurrence dependence. On a higher syntactic level of integration, however, no more could be said than that unit A presupposes unit B *for its function* in the next higher unit serving as frame; this I have called function dependence.

Dependences cannot be ascertained by a simple inspection technique; they require a testing procedure using informants or extensive textual materials.

The procedure is what I have called dropping: that is, in an utterance containing both A and B (or both A, B, and C), omit one of the units and inspect the resultant truncated utterance.

Two observations are pertinent to this examination. One is that the multiword examples very frequently contain a single noun or verb, accompanied by a non-inflected word within the same informant word. A second observation stems from the comparison of these multiword examples to one-word examples: verbs and nouns frequently constitute one-word informant words, whereas the above-mentioned accompanying words do not. An assumption of unilateral occurrence dependence of the accompanying words on nouns or verbs thus becomes reasonable.

The testing step now consists in subjecting all the multiword examples described above to the dropping test (see footnote 4). The test reveals that for most of these examples, the accompanying words are droppable, the nouns or verbs are not. This attests a relation of unilateral occurrence dependence in the data at hand, which can serve as the tentative defining criterion for a type of lowest-order syntactic units: true phrases. The independent words within a true phrase are named heads, the dependent words, modifiers. True phrases, the heads of which are nouns, constitute noun phrases; true phrases, the heads of which are verbs, constitute verb phrases.

The cross-checking step consists in applying this attested relation and the corresponding test to additional data, both additional elicited examples and running text. The result of this cross-check is to validate the relation of unilateral occurrence dependence for all those examples and textual segments that show an internal structure equivalent to the examples on which the original tests were based. All these are now validated as true phrases. All remaining examples and textual segments constitute the legitimately unaccounted residue.

In the course of the cross-check, further regularities are detected which become the subject matter of additional analytic cycles. In these cycles, the now validated relation of unilateral dependence can be introduced as a

---

For occurrence dependence, the dropping test will work as follows: A is dependent on B in an utterance containing both, if an otherwise identical utterance, but from which A is dropped, is also occurrent in the text, or is accepted by the informant as viable, B is dependent on A if the utterance from which A is dropped is non-occurrent in the text, or is not accepted by the informant as viable.

For function dependence, the dropping test will work as follows: A is dependent on B in an utterance containing both, if in an otherwise identical utterance, but from which A is dropped, the function of B remains unaltered. B is dependent on A if in the utterance from which A is dropped the function of B changes' (Garvin, 1958, pp. 628–9).

criterion for rightness in applying a substitution technique[5] as an operational test: a substitution will be considered right if the same relation of unilateral dependence is present after the substitution as before. This leads to the definition of additional varieties of true phrases by right substitutability of certain word sequences for the heads or modifiers of true phrases as previously defined; it also leads to the definition of units below the phrase level, called expressions, constituted by the right substitutes for heads and modifiers. Finally, this series of cycles includes the detection of a new relation of mutual dependence which, after proper attestation and validation, becomes the defining criterion for an additional phrase type.

The final cross-check in the last of this series of cycles suggests that, in spite of the definition of two types of phrases, the analysis of the phrase level is not complete. The residue contains a considerable number of putative units which impressionistically appear to have the same function in a larger context as phrases, but do not meet the defining criteria of unilateral or mutual occurrence dependence of words or word sequences within the putative units.

The impression of functional equivalence suggests right substitutability (see footnote 5) for phrases as a likely defining criterion for these putative units. In the search for a relation which is to serve as the criterion for the rightness of these substitutions, no purely distributional property is found.

---

5. 'Substitution consists in replacing a portion of an utterance by, in syntax, some other word or word sequence, and in comparing the result of the substitution to the original utterance. The purpose of the substitution technique is often to arrive at classes: sequences substitutable for each other are said to constitute a class. . . . The crux of the substitution technique is the extent to which substitutions are to be permitted, and the criteria by which certain substitutions are to be accepted as right and others rejected as wrong.

It is often stated that, for a substitution to be right, the result of the substitution must be a viable utterance. This is obviously correct; it is equally obviously insufficient. . . .

Impressionistically, those substitutions seem more right than wrong which yield utterances that are not only viable, but also in some manner similar to the original utterance. The next step in explicating the substitution technique is to define this similarity more closely.

Again impressionistically, this similarity between the original utterance and the result of a substitution admitted as right is one of grammatical meaning. A more rigorous analysis of this similarity of grammatical meaning will reveal that it in turn is the resultant of a maintained continuity of structural relations before and after the substitution. Thus, substitution can be defined as right if the result of the substitution is not merely a viable utterance, but an utterance containing the same structural relations of a defined kind as the original utterance. These relations will be of agreement, government and/or dependence, as the case may be for the given language.

This continuity of structural relations will then serve as a control over the use of the substitution technique' (Garvin, 1958, pp. 631–2).

IV

The above observation forces us to re-examine the previously held basic assumption that, once a form-meaning technique has been used to ascertain morphs and morphemes, the remaining morphemic analysis can be conducted by distributional techniques only.[6] In order to proceed beyond phrases, it becomes necessary to admit the existence of a separate category of form-meaning correspondences, different from those underlying morphs and morphemes, as primary characteristics of linguistic structure.

These additional primary givens are the clause functions, that is, the relational meanings (such as the subject or predicate meaning) carried by phrase-level units within the clause.[7]

The clause functions are defined in terms of form-meaning correspondences, just as are morphs and morphemes. But whereas in morphs and morphemes a one-to-one correspondence of forms to meanings is approximated, in the case of the clause functions there is a correspondence of an unrestricted set of alternative forms to each of a very restricted number of meanings: the interconnected meanings implied by the traditional designations of the parts of a sentence, such as 'subject', 'predicate', 'object', are carried by units belonging to certain morphemic classes of unrestricted membership (such as word classes or phrase types). The overall formal characteristics of these morphemic classes, in turn, provide the basic recognition criteria by which the clause functions can be observed. Thus, a word class of verbs and a phrase type of verbal phrases allow the inference of a predicate function. A word class of nouns and a phrase type of nominal phrases allow the inference of a nominal clause function, that is, one of the subject or object functions. A nominal morphemic class, however, is by itself not a sufficient definite recognition criterion to indicate which of several possible nominal clause functions is present. More detailed recognition criteria of varying degrees of definiteness may in a given language be provided by particular inflectional morphemes (such as the case endings of nouns or various indicator morphemes in the verb), or by the sequential order of the phrase-level units.

It was stated above that the clause functions are interconnected. Their interconnections are relations of presupposition, of the kind illustrated by the observation that the subject and object meanings are differentiated only in the presence of a predicate meaning, but not conversely. These are the function dependences referred to in footnote 4.

---

[6] See my previous view that syntactic analysis can be conducted simply by 'using as our criteria a set of dependences and substitutions' (Garvin, 1954, p. 348).

[7] Note that the clause, rather than the sentence, is the unit within which these meaningful relations are recognized.

The domain of the clause functions does not extend over the whole of the clause, but only over the functionally interconnected phrase-level units (the clause members) that are the carriers of the functions. The clause members make up the body of the clause. Such additional words and word sequences as may be present in the clause but carry no clause function (as for instance, certain particles) are not included in the clause members but form a separate category in terms of their external functioning: they are clause dependents. The clause functions thus allow a differentiation of the constituents of a clause into clause members and clause dependents.

The operational tests of linguistic analysis, such as the previously mentioned dropping and inversion tests, are based on distributional properties. Since the clause functions are not distributional features, but form-meaning correspondences, they are therefore not amenable to such operational tests. Order as a recognition criterion of the clause functions, as well as the function dependences, on the other hand, are distributional corollaries of the clause-function relation. They are therefore susceptible to operational tests: a permutation test can be applied to order, a dropping test can be applied to the function dependences.

The permutation test consists in rearranging the order of the nominal members of a clause with respect to the predicate and observing how this changed order affects the subject and object functions of nominal clause members: if different clause functions are assumed by the same morphemic unit as a result of changed order, order as a recognition criterion of clause function can be considered attested.

In applying the dropping test for function dependences, it will be discovered that in the absence of a predicate there occurs only a syncretism of the subject and object functions, which can be called the syncretic function (Hjelmslev, 1953, p. 56).

The above operational tests deal merely with the corollaries of the clause functions, and not with the clause functions themselves. The inapplicability of operational tests to the clause functions does not, however, mean that the latter cannot be subjected to some other suitable procedural controls. Being form-meaning correspondences, the clause functions are amenable to control by elicitation. Just as the form-meaning correspondences underlying morphs and morphemes serve as the basis for the elicitation of morphological paradigms, so the clause functions can serve as the basis for the elicitation of syntactic paradigms.

The elicitation of paradigms is a form-meaning technique in which meanings are controlled as the independent variable, in order to observe corresponding forms as the dependent variable.[8] It is based on certain tacit assumptions about linguistic universals, which can be rendered explicit as follows:

---

[8] My statement in a previous paper is pertinent to this discussion: 'In linguistic field work, meaning is used as the independent variable to elicit forms. This is done by asking

The assumption underlying morphological elicitation is that certain grammatical meanings have a high probability of being in a near one-to-one correspondence to sets of forms in any given language. Thus, it is expected that some set of person markers will correspond to the usual set of 'person' meanings. Sometimes the expectation of a certain form-meaning correspondence is limited to a particular language family or area, based on previous formal or informal knowledge. An example of this is the assumption that in a Malayo-Polynesian language, there is a great likelihood of a formal difference corresponding to the meaning difference 'first person plural inclusive' (i.e., we = I, you, and possibly others) as opposed to 'first person plural exclusive' (i.e., we = I and others, but not you).

The paradigms elicited on the basis of these assumptions provide the necessary starting point for the analysis. They constitute the framework for the subsequent observation and manipulation of the conditions that can not be foreseen in the very general initial assumptions.

Syntactic elicitation is based on the assumption of the co-variance of morphemic class membership with clause function. It is aided by the prior observation of the formal recognition markers of the clause functions, that is, in addition to morphemic class membership, the particular morphemes and features of order discussed further above.

Instead of informant words, as was the case in morphological elicitation, there will now be elicited informant sentences, that is, utterances which impressionistically constitute single-clause sentences. The clause functions will be controlled as the independent variable, and resulting conditions of morphology and phrase syntax will be observed as the dependent variable. In order to conduct this observation in terms of the conditions assumed to be relevant to the exercise of the clause functions, examples for elicitation will be selected on the basis of the anticipated morphemic composition of the desired informant responses, with the aim of obtaining informant sentences of the appropriate structure and of excluding any words not forming part of the members of relevant classes.

The role of syntactic elicitation in the analytic process is thus different from that of morphological elicitation. Whereas morphological elicitation serves to provide the data base for the initial cycle of morphemic analysis,

---

the informant to translate controlled sets of sequences of known meaning. These sequences are held semantically constant, with the exception of one minimal controlled variation, on the assumption that the co-variant in the informant's responses of this minimal semantic variation may be a minimal formal variation. If such a co-variance results, the particular line of questioning is considered unproductive and the assumption of a paradigmatic set for which these questions were designed is abandoned' (Garvin, 1959, p. 145).

syntactic elicitation serves as a check on insights gained from the inspection of a previously acquired data base. Though not an operational test in the sense discussed further above, syntactic elicitation therefore will be included in the testing step of an appropriate syntactic cycle.

V

In the analysis of Ponapean, the analytic objective of the cycle following the definition of phrases can now be formulated as the specification of the clause functions particular to that language.

The data base for this new cycle is formed by examples of appropriate morphemic composition selected from the morphological field records and from running text. The relevant morphemic classes are the morphologically defined word classes of nouns, verbs, and locative demonstratives (the latter correspond very roughly to Indo-European adverbs of time and space such as *here*, *now*, *there*, *then*) and the phrase types that have been defined in previous syntactic cycles, namely noun phrases and other nominal phrases, verb phrases and other verbal phrases. Putative clause dependents are ignored in the selection and subsequent examination of the examples.

In the examination step, five clause functions can be ascertained for Ponapean: a subject function, a predicate function, a direct object function, an indirect object function, and a complement function. A function dependence of subject and objects on the predicate, as well as the functional independence of the complement, are observed. The functionally interdependent clause members (subject, predicate, objects) are termed the core members of the clause, the complement is termed a marginal clause member. The presence of certain indicator morphemes in the verb is noted whenever an indirect object is present. In addition, an apparently fixed sequential order is observed among the core members of the clause: subject, predicate, direct object, indirect object. There appear to be no order restrictions with respect to marginal clause members, that is, complements. Finally, the following associations of clause function and morphemic class membership are registered: nouns and nominal phrases are found to assume the subject function and the two object functions (subjects and objects are therefore called nominal clause members), verbs and verbal phrases are found to assume the predicate function, locative demonstratives are found to assume the complement function.

In the testing step, syntactic elicitation is used to attest the association of clause function and morphemic class membership. The dropping test is used to attest function dependences, the permutation test is used to attest the order relation.

The cross-checking step consists in applying the attested relations and corresponding tests to additional data (including examples culled from running text), consisting, as before, of combinations of previously defined phrase-level units. The result of this cross-check is to validate the function dependences, but not the relation of order to clause function: a few examples are discovered in which, instead of the expected order, an object is found to precede the predicate, or a subject is found to follow it. In these examples, the lexical meaning of the nominal clause members in question is such that only one clause function is culturally acceptable, irrespective of position with respect to the predicate. The relation of order to clause function is therefore found to be valid only conditionally, and can be superseded under appropriate lexical conditions. This allows the conclusion that order is required as a recognition criterion for clause function only if the lexical meaning of the morphemic unit in question is ambiguous in this respect.

The independent character of the clause functions is underlined by the discovery of additional associations of clause function with morphemic class, as examples are culled from text: nouns and nominal phrases are found to function not only as subject and objects, but also as predicate and complements. Verbs and verbal phrases, however, continue to be restricted to the predicate function, locative demonstratives to the complement function.

In a number of the examples culled from text, moreover, it is found that more than one member of a relevant morphemic class, linked by certain particles (comparable to conjunctions), can share the same single clause function. A related observation pertains to the number of clause members present: while in any given example, more than one complement may be found, no example contains more than one each of the core members of the clause. It can thus be assumed that there may be only one subject, predicate, direct object, indirect object, per clause, but that there may be more than one complement.

VI

It is now possible to return to the earlier objective (see the end of Section III) of defining the additional units of the phrase level. The conditions for a substitution to define these units can now be stated. The substitution frame will be the body of the clause, consisting of a maximum of one subject, predicate, direct object, indirect object, and one or more complements. The criterion for the rightness of the substitutions will be the preservation of the same clause functions after the substitution as before. The units sought are to be defined as substitutes for nouns or nominal phrases, verbs or verbal phrases, locative demonstratives. Units rightly substitutable for nouns or nominal phrases will be termed nominal phrase substitutes; units rightly substitutable for verbs or verbal phrases will be termed verbal phrase substitutes; units rightly substitutable for locative demonstratives will be termed locatives.

These substitutions are carried out in a new series of cycles, for which the data base is taken from running text. In the examination step, putative clauses are selected. In the testing step, the substitution technique is applied. The cross-check consists in the examination of additional text.

This latest series of cycles completes the analysis of the body of the clause. Attention can now be given to the clause dependents which so far have been deliberately ignored.

The examination of putative clauses suggests that the clause dependents are in various relations of occurrence dependence on the body of the clause as a whole, or on an entire additional clause. These occurrence dependences and their role in defining the syntactic function of clause dependents become the subject of a further series of cycles, the ultimate result of which is the proper definition of the clause. A final series of cycles allows the refinement of these and further dependences and leads to the definition of multi-clause sentences and of the relations of clause coordination and subordination. This completes the syntactic description.

VII

A significant general conclusion can be drawn from the above discussion of inductive procedure as applied to a particular language.

It is reasonable to assume that, since the clause functions appear to be primary characteristics of linguistic structure, a purely distributional analysis will not account for all intuitively observable syntactic relations. It will have to be coupled with the use of a form-meaning technique. The major problem in the analysis of any given language will then be the determination of the point at which distributional procedures are exhausted, and only the introduction of a form-meaning cycle will allow the continued use of distributional techniques to complete the analysis. It is procedurally significant to note that the form-meaning cycle is introduced only when it becomes necessary, and only to the extent that it is required to allow further distributional analysis.

The common feature in the application of these techniques is the use of informally specified units (behavioral units obtained by elicitation, or putative units inferred from the examination of data) as initial test frames for the discovery of the rigorous criteria on which a procedurally proper description is based.

References

GARVIN, P. L. (1954), 'Delimitation of syntactic units', *Language*, vol. 30, pp. 345–8.

GARVIN, P. L. (1957), 'On the relative tractability of morphological data', *Word*, vol. 13, pp. 12–23.

GARVIN, P. L. (1958), 'Syntactic units and operations', *Proc. Eighth Internat. Congress Linguists*, Oslo, pp. 626–32.

GARVIN, P. L. (1959), *On Structuralist Method*, Georgetown University Monograph Series on Language and Linguistics, no. 11, pp. 145–8; 167–70.

GARVIN, P. L. (1963), 'The definitional model of language', in P. L. Garvin (ed.), *Natural Language and the Computer*, McGraw-Hill.

GARVIN, P. L., and MATHIOT, M. (1958), 'Fused units in prosodic analysis', *Word*, vol. 14, pp. 178–86.

HJELMSLEV, L. (1953), *Prolegomena to a Theory of Language*, tr. F. Whitfield, Johns Hopkins Press.

22 K. Bergsland

Morphological Analysis and Syntactic Reconstruction in Eskimo-Aleut

K. Bergsland, 'Morphological analysis and syntactic reconstruction in Eskimo-Aleut', *Proceedings of the Ninth International Congress (1962) of Linguists*, 1964, pp. 1009–15.

1. In Eskimo, as is well-known, there are two main types of verbal constructions, 1. intransitive: verbal form marked for one grammatical person (the subject), in number agreement with a noun in the so-called absolutive case (sg. zero. du. *-k*, pl. *-t*), e.g. Greenlandic indic. *teriangniaq pangalig-poq* 'the fox ran', *teriangnia-t pangalig-pu-t* 'the foxes ran'; 2. transitive: verbal form marked for two grammatical persons, the one (the object) in agreement with a noun in the absolutive case, the other (the actor) in agreement with a noun in the so-called relative case (sg. *-m*, Grl. *-p*, du. and pl. = abs.), e.g. *teriangnia-p orssoq neri-vâ /-va-a/* 'the fox ate the blubber', *teriangnia-t orssoq neri-vât /-va-a-t/* 'the foxes ate the blubber', *teriangnia-p nerdleri-t nererêrsima-va-i* 'the fox had already eaten the geese'.

Verbs that occur in both constructions have, in the intransitive one, (a) an active sense, e.g. *neri-voq* 'he eats, ate', or (b) a reflexive sense, e.g. *toqúpoq* 'he killed himself' vs. transitive *toqúpâ* 'he killed him', passive *toqutauvoq* 'he was killed' (passive participle+*-u-* 'be') and *toqúneqarpoq* id. (verbal noun+*-qar-* 'have'). From the latter, and from other transitive verbs, are derived so-called half-transitive verbs, active intransitive verbs which may have a complement in the instrumental case corresponding to the object of the underlying transitive verb, e.g. intr. part. *inung-mik toqut-si-ssoq* 'who kills or has killed a person, murderer' (Kleinschmidt, 1851, pp. 54 ff.).

From the point of view of the nominal cases, and with special regard to the verbs of the semantic type (b), one may say that the transitive construction is passive (Thalbitzer, 1911, pp. 1058 ff.). From a syntactically more comprehensive point of view it appears as active: a so-called 4., i.e. reflexive 3., person suffix refers to a subject/actor, Kleinschmidt's 'projekt' (1851, p. 15), rather than to an absolutive term (subject/object), e.g. N.Alas. indic. *arna-t īnī-mīg-nun* (sg/pl. 4. p. pl.+allative) *aillak-tu-t* 'the women went home to their place', *aapa-ŋa-n* (sg.3.p.sg.+rel.) *uqullautī-ga-a irni* (sg. *irnīq*+4.p.sg. *-ni*) 'the ('his', the son's) father said to his son'; W.Esk. Kuskokwim *qayâ-ne ayautâ* 'he takes away (or removes) his kayak', half-tr. *qayâ-mi-nik ayautsioq* 'he removes his kayak' (Hinz 1944, 43). In respect to alternative interpretations the transitive construction of Eskimo recalls the ergative construction of Georgian as discussed by Vogt (1950).

2. To a nominal subject/object in the absolutive case correspond subject/object suffixes for the 1., 2. and 4. persons (Hammerich 1936, pp. 159 ff.), e.g. Grl. *akivu-nga /-ŋa/* 'I answer', *akivavsi-nga* 'you (pl.) answer me'. In the contemporative mode, where the subject/actor usually is the same as that of the verb to which the contemporative is joined, there is normally one person suffix only, in either function, e.g. N. Alas. . . . *mīkī-plu-ni* (4.p.sg.) *inuiyuminait-lu-tīŋ* (4.p.pl.) *īluqarmīŋ isagutī-kp-a-tīŋ* (conditional 3.p.sg.+4.p.pl.) '(they were afraid of him) he being small, being able to kill them all, if he attacked them'. The corresponding 3.p. suffixes, sg. *-gu*, du. *-gik*, pl.*-git*, WE *-ku, -kīk, -ki*, which occur also in other modes, are object suffixes only.

In the indicative and the participial, the suffixes for object 3.p+actor 1-4.p. are the same as the suffixes found in nouns, in the absolutive case, for number+1-4.p. possessor (Kleinschmidt, 1851, p. 51), cf. N. Alas. *tuttu-t ayulirī-t-ka* 'I am missing the caribou (pl.)' and *qītunra-t-ka* 'my children'; *arna-t qinīra-a-t* 'the women looked at her' and *umialīkpai-t nīqa-a-t* 'the rich people's food'. With a 1., 2. and 4.p. possessor/actor the markers for numerus possessi/objecti can be identified as variants of the simple number suffixes, sg. zero, du. *-k* and pl. *-t*, which in intransitive forms mark 3.p. subject and so correspond to the subject/object suffixes for the other persons. With a 3.p. possessor/actor the suffixes for numerus possessi/objecti are sg. *-a- /-ŋa-*, du. *-k-*, pl. *-i-/-ŋi-*, followed by the simple number suffixes which here agree with the possessor/actor in the relative case (Hammerich, 1936, p. 203).

From the formal identity of these verbal and nominal constructions one has concluded that the verbal forms have a nominal nature (Thalbitzer, 1911, pp. 1057 ff.) and become predicates only 'durch das hinzutreten des existenzprädikats' and that there is in '*panínguit asavara* "dein töchterchen [ist] mein geliebtes/meine liebe" genau dasselbe subject-prädikatverhältnis wie in *panínguit tikípoq* "dein töchterchen kommt" ' (Hammerich, 1936, pp. 210 ff.; cf. Sauvageot, 1953). However, in the former sentence an additional 'her/she alone', for example, would have to be in the 3. person (*kisiat*), whereas in the latter it would have to be in the 4. person (*kisime*), so the interpretation is reconstructive rather than descriptive. What is a participle in one dialect may be an indicative in another, but in general an Eskimo noun, unless verbalized with a derivational suffix (e.g. *-u-* 'be'), does not constitute a predicate in the way of a verb (Kleinschmidt 1851, pp. 67 ff.; Hammerich, 1936, p. 212).

3. The actor markers that precede the object suffixes in the dependent modes – causative (conjunctive, relative preterit) etc. and conditional (sub-

junctive, relative future) – can be identified as modified forms of the subject markers in the same modes. (With some exceptions the same holds good for the interrogative and the imperative-optative.) These subject/actor suffixes can be further identified with the suffixes of nouns in the relative sg.+1-4.p. possessor (Kleinschmidt, 1851, p. 50). Cf. N. Alas. *paqīt-ku-ma* 'when I find', *naat-ku-p-kit* (WE-*m-ki*) 'when I finish them' and *nunama* '(of) my land'. Before the object suffixes, as before local case suffixes, most of the 3.p. forms, which have special final alternants of the relative suffix, look like absolutive forms (Hammerich, 1936, p. 203), e.g. *-a-ŋa* 'he me' vs. intr., sg.3.p.sg. rel. *-a-n* (Grl. *-at*); WE, N. Alas. *-at-ŋa* (Grl. *-aŋŋa*) 'they me' vs. intr., 3.p.pl. rel. *-at-a*. This can be accounted for in terms of assimilation and syncope, cf. 1.p.sg. *-p-*/*-m-* and *-ma* above, and, for example, the different treatment of *t*+*ŋ* in WE Kusk. /*tīkiz-ŋaitut*/ 'they will not arrive' (*tīkit-*), N. Alas. *min-ŋaujaqtuaq* 'which has alighted' (*mit-*).

Unlike nouns in the relative case, the verbal relative forms have no correlate in a superordinate term with a 3.p. possessor/actor suffix, cf. N. Alas. *aŋugaatsia-m īglu-a-n kīlu-a-ni* (loc.) 'in the rear of the elderly man's house' and *iqsinaqtaaq īli-kp-a-n irni aanruaqarnira-a* 'in case that a monster turn up, she gave her son an amulet' (the final *-a* refers to *irni*, not to *ilikpan*). To reduce the verbal construction to the nominal one, one has assumed ellipsis of a term corresponding to *kīlu-a-ni* in our nominal construction (Thalbitzer, 1930, p. 327) or else ellipsis of a more 'abstract' nature (Hammerich, 1936, pp. 217 ff.). Actually, the constructions differ also in respect to the other term of the relation: with a nominal term like *īglu-a-n* goes a nominal term in the relative case (the possessor *aŋugaatsia-m*), with an intransitive verb form like *īlikp-a-n* goes a nominal term in the absolutive case (the subject *iqsinaqtaaq*). Thus, the 'same' suffix complex *-a-n* enters into quite different syntactical constructions.

4. In the indicative and the participial, one finds before the 1., 2. and 4.p. object suffixes the same actor markers as in the dependent modes i.e. relative forms (except that in the WE 3.p. du = pl. actor+du. and pl. object the object suffixes seem to be added to the 3.p.pl.+3.p.pl., e.g. Kusk. indic. *-it-kut*, depend. *-at-kut* 'they us', cf. indic. and depend. *-a-kut* 'he us'; EE *-atīgut* 'he/they us', which like WE *-at-kut* seems to reflect 3.p.pl.rel. **-ata+kut*, belongs to both sets of modes).

These forms can hardly be explained by the fact that they are transitive (Hammerich, 1936, p. 201) – transitive verb forms have a nominal actor in the relative case, but here the verb itself has the shape of a relative form. Actually, apart from the intransitive participle, there are only faint traces of a case distinction in the Eskimo verb. In the conditional the 4.p. forms mostly have the absolutive shape (sg. *-nī*, du. *-nīk* and pl. *-nīŋ* with *-n-* from the singular, Siberia *-yīk*, *-yīŋ*) but the opposition to the other, relative, forms is one of person rather than one of case. In other words, 'relative' is here an exclusively morphophonemic term.

5. A descriptive analysis of Eskimo that works down from the higher levels yields structural units (e.g. object 3.p. sg.) which cannot be profitably isolated by any uniform method of morphophonemic cutting but can easily be described in terms of overlapping morphophonemic units (e.g. *-a* 'he him/his one'). It is reasonable to assume that the overlapping is the result of a historical process, and by assigning uniform meanings and functions to the morphophonemic units one can no doubt build up some other structure behind the actual one. However, if the actual structure is assumed to be the outcome of a unique process that would constitute the only possible 'explanation' of it, how could one select the diagnostic meanings and functions? If it is assumed to be of a TYPE that can have only one type of antecedent, the descriptive analysis would seem to be sufficient. And if one admits the possibility of more than one type of antecedent, how could one determine which one it actually was? A tentative syntactical reconstruction seems to be of interest only for the comparison with one or more different languages assumed, on the basis of specific correspondences on the morphemic level, to reflect the same proto-language. In our case, then, what can the comparison with Aleut tell us about the structure of Proto-Eskimo?

6. In Aleut there are object suffixes only for the 3. person, sg. *-ka*, du. *-ki-g*, pl. *-ki-n*, Atka *-ki-s* (*-n*/*-s* = Esk. *-t*), in the contemporative, the imperative and in a dependent mode (cf. 2.). For the other persons there are personal pronouns, independent or proclitic as object, enclitic as subject, e.g. *tiŋ ukurtanar-tgici* 'you (pl.) have seen me' (without *tgici*: 'he has seen me'). These pronouns, which consist of a stem *tgi-*/*ti-*+possessive suffixes, probably are the cognates of the Eskimo subject/object suffixes; 2.p.sg. *tgin* accounts for both alternants of the Eskimo suffix *-tīn*/*kīn-* (Hammerich, 1936, p. 172) and the comparison can easily be justified phonologically also for the rest of the set. Of the Eskimo independent pronouns, some contain the respective subject/object suffixes (1.p.sg. *uva-ŋa* etc.), the other ones have another set of Aleut cognates (2.p.sg. *īl-vit* etc., Aleut *i-min*, Atka *i-mis* 'to, for you', etc.).

If from these specific morphemic correspondences one may infer that the independent pronouns have become object suffixes in Eskimo, and that they have become so on the analogy of the 3.p. object suffixes, a simple explanation of the enigmatic relative forms in the Eskimo indicative presents itself: together with the new object suffixes the relative suffix could have been transferred from dependent modes. If the pronouns have become object suffixes in part by directly joining a preceding subordinate verb form, the displacement of the pronoun could also have turned the Aleut type *tgin asratikur* 'he killed himself' into the above-mentioned Eskimo semantic type (b) *toqúpoq*, cf. Atka *tayarur anrarinas tunurta: saku: tgin haqatalaga: ran artakur* 'the man / the people / [whom, while] talked about / himself / did not know, i.e. the man did not know that the people were talking about him'. Quite possibly, then, the Eskimo opposition intransitive vs. transitive (1.) reflects an opposition of the type found in Aleut.

7. As first realized by Waldemar Jochelson (1912, pp. 1044 ff.), the Aleut forms with actor/possessor suffixes differ from the morphologically corresponding Eskimo forms in two important syntactical respects, exemplified by Jochelson as follows: 1. *anrari-r* (abs.sg.) *qa-r* (do.) *su-ku-r* (indic. 3.p.sg.) 'the person took the/a fish', with fully explicit object (*qa-r* and no suffixal reference to it in the verb, vs. *anrari-m* (rel. sg.) *su-ku-*: 'the person took it', with suffixal reference (*-:* = post-consonantic *-a*, Esk. *-a*) to the implicit object, and *qa-r anrari-m su-ku-:* 'the fish, the person took it', with reference to the absolutive term preceding the actor. 2. *ula-m il-a-n* (3.p.sg.loc.) *uŋuci-ku-qin* (= *-ku-r+tiŋ*) 'I am sitting in the house', with fully explicit local complement, vs. *il-a-n uŋuci-ku-ŋ* 'I am sitting in it', with suffixal reference (*-a-* and *-Ø-ŋ*) to the implicit part of the complement. For the last point cf. Atka *ukina-r* (explicit object) *ila-ŋi-:n* (3.p.pl.ablat.) *agati+arta-ku-ni-ŋ* 'I took the knife away from them (*-ŋi-*, *-ni-*)'.

To a term with a 4.p. suffix, which is fully explicit in the indicated sense, there can be no suffixal reference in the final verb, cf. Atka *fuřa:ski-:n curta-ku-r hiŋa* 'he has his cap on, that [person]'. Just as here the demonstrative stem form *hiŋa* points to an 'implicit' subject, it points to implicit terms of the two other kinds in *su-ku-ŋ hiŋa* 'I took it, that [object]', *asa-: haqata-lakari-ŋ hama* 'his name I do not know, that [person's]'. So far one is free to interprete the type *qa-r anrari-m su-ku-:* as 'passive' (although there are also derivational passive forms, with removal of the actor, e.g. *qa-r su-lga-ku-r* 'the fish was taken'). However, if *qa-r*, the absolutive term to which there is suffixal reference, is the grammatical subject here, the same holds good, not only for *anrari-r* in *anrari-r qa-r su-ku-r*, but also for *itraygi-r* in the type *itraygi-r cŋa-: usa-ku-r* 'the reindeer, its hair falls off, i.e. the reindeer sheds its hair' (not the same as **itraygi-m c.u.* 'the reindeer's h.f.o.'), so the predicates are CLAUSES.

8. The Aleut indicative, which seems to be the cognate of the Greenlandic transitive participle and the usual indicative in N. Alaskan, does not differ morphologically from a noun and so fulfills the morphological conditions for interpreting the verbal constructions in question as 'nominal', but it

does not fulfill the syntactical expectations that are implicit in this interpretation, cf. Atka *kimi-ku-r ukurtal aŋali-ŋ* 'it [the airplane] descending / I saw it today', *aniqdu-m su-ku-:n uqidusadu:ka-lakar-a* 'the child / what it takes / it will not return it'. The suffixal reference in the final verbs (*-Ø-ŋ* and *-a*) shows that *kimi-ku-r* and *su-ku:n* have one more 'slot' than the nominal objects *qa-r* and *fuřa-ski-:n* in *anrari-r qa-r su-ku-r* and *fuřa: ski-:n curta-ku-r hiŋa* above. When the slot is filled, for example with an explicit 'airplane', there is no reference to it in the final verb.

In Aleut, as in Eskimo, the indicative constitutes a clause by itself, whereas a noun normally needs 'das hinzuterten des existenzprädikats', viz. the verb *a-* 'be'.

9. The Aleut participles (agentive *-na-*, cf. Esk. verbal noun *-nīq*; passive *-qa-/-ka-*, WE *-kaq*, EE *-gaq*; indefinite -Ø-) share with the indicative the function of a final predicate but differ from it in other syntactical respects, cf. *ayugi:rin+arta-na-s ukurta-na-r-t*[*gin*] *i:* 'did you see some going out?', *igarta-r haqa-na-: ukurta-na-r-t*[*gin*] *i:* 'did you see the airplane coming?'. Both are full sentences, as shown by the final predicate (*-na-r+tgin*). In the former, the participle (*-nas-s* pl.) alone fills the 'object slot' of the transitive verb *úkurta-* and so turns out to have no subject slot itself. In the latter, the 3.p. suffix(-:) adds that slot and refers to the absolutive term *igarta-r* (incidentally an indefinite participle, 'flyer'). In this sense also the conditional, which seems to be the cognate of the Eskimo conditional, is a participle, cf. *igarta-r ama:nu-gu-:* 'when the airplane leaves'. Here, as in the Eskimo intransitive dependent verb forms, the 3.p. 'possessive' suffix refers to an absolutive, non-subordinate, term. In other phrases it refers to a term in the relative case, e.g. *aŋali-m haqa-:* 'the coming of the daylight, the coming daylight', and, with one more slot, *anrari-m su-ku-:* 'the person took it'. The reference is the 'same', the constructions are different.

The Aleut conditional shows case distinction, e.g. 4.p.sg. abs. *sara:ran aqali-gu-:n* 'whenever he was about to sleep (he at the same time . . .)', rel. *tgin saraniqada-gu-:m tgin guŋtukalizar* 'when he has fallen asleep, he usually starts snoring'. Here, as in the case of the Eskimo dependent modes, there is no term with 3.p. reference to the relative conditional. In this respect the conditional differs from a noun and goes with the indicative, cf. *anqarta-ku-m* (rel.sg.) *haqa:ran+arikur* 'he went away but will come back in a moment'. The relative case may be said to mark subordination but it too can enter into different syntactical constructions.

10. If the comparison with Aleut indicates a possible structural background for the morphological pecularities of the Eskimo verb, it does not explain all of its constructions. The Eskimo half-transitive verbs (1.) are comparable to Aleut intransitive derivatives such as *asrat-gari-*(*laga-da* 'thou shalt not) kill', but the instrumental complement of an Eskimo half-

transitive verb has no other analogue in Aleut than the absolutive object of an ordinary transitive verb. The absolutive object is in keeping with the absolutive pronouns which in Eskimo have become object suffixes (6.) but in Aleut there is no instrumental case at all. The morphological analysis of the Aleut case forms shows that the system of local cases has been reduced, so it is reasonable to assume that the instrumental has been lost, but to determine its exact functions in the proto-language and the causes and effects of its loss in Aleut is hardly possible. Starting with this aspect of the comparative problem one could build a different model of the proto-language, and it is evident that the comparative method does not permit a syntactical reconstruction in any real, historical, sense.

The comparison of Eskimo and Aleut demonstrates the impossibility of determining the syntactical structure from the morphemes and their meanings. Without a syntactical frame of reference, however, comparisons of grammatical elements are vague and can claim no great interest. Our case shows that the syntactical frame can not be taken for granted but has to be established by comparison. In other words, syntactical reconstruction is equivalent to stating correspondence formulae on the higher levels which can make the comparisons on the morphemic level more meaningful. By comparison one can get no additional material for linguistic typology but one can gain insight into how linguistic structures are interrelated and to what extent the common elements are determined by their own setting.

References

Bergsland, K. (1959), 'Aleut dialects of Atka and Attu', *Trans. Amer. philosoph. Soc.*, vol. 49, p. 3.

Hammerich, L. L. (1936), 'Personalendungen und Verbalsystem in Eskimoischen', *Det Kgl. Danske Videnskabernes Selskab. Historisk-filologiske Meddelelser*, vol. 23, p. 2, Copenhagen.

Hinz, Rev. J. (1944), *Grammar and Vocabulary of the Eskimo Language*, Moravian Church, Bethlehem, Pennsylvania.

Ingstad, H., texts recorded on tape in Brooks Range, to be edited by author.

Jochelson, W. (1912), 'Zametki o foneticheskikh i strukturnykh osnovakh aleutskago yazyka', *Izvestiya Rossiyskoy Akademii Nauk*, pp. 1031–46.

Kleinschmidt, S. (1851), *Grammatik der grönländischen Sprache*, Berlin.

Menovshchikov, G. A. (1954), *Eskimossko-russkiy slovar*, Leningrad.

Menovshchikov, G. A. (1960), *Eskimosskiy yazyk*, Leningrad.

Sauvageot, A. (1953), 'Caractere ouraloïde du verbe eskimo', *Bull. de la Soc. de Linguistique de Paris*, vol. 49, pp. 107–21.

Thalbitzer, W. (1911), 'Eskimo. Handbook of American Indian Languages' *Bureau Amer. Ethnol., Bull.* 40, pp. 967–1069.

Thalbitzer, W. (1930), 'The absolutive and the relative in Eskimo', in *A Grammatical Miscellany Offered to Otto Jespersen on his Seventieth Birthday* pp. 319–29, Copenhagen.

Vogt, H. (1950), 'Un aspect du probleme actif-passif dans le verbe' *J. Psychologie*, pp. 130–38.

23 B. T. Tervoort

You Me Downtown Movie Fun?

B. T. Tervoort, 'You me downtown movie fun?', *Lingua*, vol. 21, 1968, pp. 455–65.

Every disciple of Anton Reichling will remember the sequence *wit huis op berg* and its twenty-four permutations; he used it to demonstrate syntactic grouping. To quote him freely: words are grouped according to principles of sequence, according to patterns of intonation, and – in some languages – according to morphological correlation.

While analysing the communication of deaf children, I found corroboration of Reichling's thesis that the principles of sequence are the most basic ones. Working this out for a system with a very weak structure and a great freedom of order might prove to be relevant to a better understanding of some linguistic fundamentals. Even in this system, enlarging a group by initial or final addition affects its coherence less than penetrating it. If an insert penetrates a group (a) it is incorporated into the group, or (b) it splits up the group. Whether (a) or (b) is effectuated, depends on several conditions which will be discussed here.

Some introductory information on the communication of deaf children among themselves will clarify the subsequent analysis. I shall avoid the term *word* and use *signum* instead, as the medium of communication can be the *word*, spoken with or without voice or fingerspelled in the air, the motivated *gesture* or the arbitrary *sign* with the hand, and, finally, the *simultaneous co-occurrence* of spoken and spelled word or of spoken word with gesture or sign. Speech + sign and sign-only are the preferred means of execution, and speech-only is very rare; apart from this, the different executions occur at random.

With these pluriform signa, the deaf sometimes produce grammatical or acceptable English sentences. If they use some speech or fingerspelling, they might mark some words morphologically; if they just sign, exoteric morphological marking is absent. Most sentences, however, are ungrammatical; I shall limit myself to these. Although they are ungrammatical English, in the private communication under discussion they are the normally accepted ones. They will be termed *esoteric*; their structure reveals the deviant system used. In this type of sentence, exoteric morphological marking is always absent, and grammatical function-words seldom occur.

Although there is some kinaesthetic motor-experience and a minimum of residual auditory perception in some cases, the communication is mainly visual, so that intonation does not function; it is replaced by mimicry and rhythmic grouping through patterns of movement. Greater reliance upon context, situation and the unavoidable visible presence of the partner is necessary to effectuate mutual understanding.

A preliminary example might clarify some of this. During a summer course for teachers of the deaf, my American students agreed that:

Who captain you me all boy?

could be interpreted in at least fourteen different ways. *Captain* was a sign with the hand to which one could add either nominal or verbal connotation, and *you, me* and *all boy* could all be (part of a) subject or direct object. Paraphrasing the sentence in either the past, the present or the future, would have multiplied the fourteen possibilities to two thousand seven hundred and forty-four. No further recurrence to syntactic ordering is possible; to solve a case of homonymy like this one, it is necessary to see the sentence executed, in order to observe the rhythmic grouping, or to have a certain amount of background information at one's disposal. The former would only be possible if the sentence was filmed; the latter conveys the fact that the utterance came from a boy who asked the question, knowing that neither he nor his partner were going to be captain of their football team. Knowing this, a correct paraphrase in grammatical English can be performed; it consists in adding function-words and morphological marking, as follows:

will be the *of* *and* *the* *s*
Who captain you, me all boy ?

During some fifteen years of studying some twenty thousand sentences of deaf children, I have noticed a very great freedom of sequence of the signa. The counterpart of this freedom, of course, is the great dependence on the situation with the visible partner as centre. I believe this freedom has to be attributed, among other things, to the basic principle of distinctive visibility in the system, as opposed to the distinctive audibility on which speech-sound sequences are built. The former principle does not imply the strict linearity of the latter but permits simultaneity, as in:

Mad you me not?

in which *mad* was signed with the right hand, *me* was only spoken, *you* consisted of the pointing left index finger, *not* was the shaking of the head, and the mimicry conveyed the question. As all five were completely simultaneous, and even prolonged together, the transcription could equally well have been:

Mad you me not?
Mad you not me?
Mad me you not?
Mad me not you?
Mad not you me?
Mad not me you?
You mad me not?
You mad not me?
You me mad not?
You me not mad?
You not mad me?
You not me mad?
Me mad you not?
Me mad not you?
Me you mad not?
Me you not mad?
Me not mad you?
Me not you mad?
Not mad you me?
Not mad me you?
Not you me mad?
Not you mad me?
Not me mad you?
Not me you mad?

which then would explain why a deaf child, wanting to speak or to write down these signa in a sequence, is at a loss about their succession. The child who uttered *Mad you me not?* was five years old; she addressed her kindergarten teacher who looked annoyed. The latter's correct interpretation of the utterance from the situation was evidenced by her answer: *No dear, I'm not mad at you*; the utterance has to be interpreted as: *You are not mad at me, are you?*

Simultaneous execution (and visual perception) of different signa evidently unifies them as one non-sequential group. Simultaneous execution, however, is the exception: more groups occur in a sequence, but it looks as if the order of the signa is irrelevant. American deaf children, for example, used the following groups of two indiscriminately:

beautiful flowers – flowers beautiful
my mother – mother my
your baby – baby your
he came – came he
track meet – meet track
fire drill – drill fire
sunny day – day sunny
we go – go we
class period – period class
basketball girl team – basketball team girl

I found the same characteristic in the communication of Dutch and Belgian deaf children, e.g.:

groot-vader – vader-groot
kap-salon – salon-kap
paarde-staart – staarte-paard
pluk-haren – haar-pluk
manne-hemd – hemd-man
doofstommen-onderwijzers – onderwijzers-doofstommen
klein meisje – meisje klein
jouw kind – kind jouw
sigaren-pijpje – pijpje-sigaar

In all samples given, the second item conveys the same as the first. All these combinations were as acceptable to them as e.g., *he said* vs *said he*, or *home bound* vs *bound home* in colloquial English, or like *meisje klein* vs *klein meisje* in poetical Dutch. There was no evidence that these children made a difference between a group of two signa (your baby) and a compound signum (firedrill). Both English and Dutch grammar, of course, have the rule that the first member of a compound word is the determining and the second the determined part which classifies the complete word: a *firedrill* is some kind of a *drill*, not a *fire*, a *paardestaart* is a *staart* (tail) not a *paard* (horse), and the word *paardestaart* is a *de*-class substantive because *staart* is, not a *het*-class substantive, like *paard*. Apparently, deaf children, at least sometimes, do not know this rule. For them, I believe, *track-meet* and *groot-vader* are as much wordgroups as *meet track, vader groot, mother my* and *meisje klein.*

To find out whether any syntactic rules of grouping could limit this freedom of order, I set up a small, unpretending '*wit huis op berg*'-experiment, whenever I was teaching summer courses in the US (five times between 1960–64) for graduate students, enrolled in teacher-training courses for the deaf. All of them were familiar with the deaf child's communication features, and some of them were deaf themselves. As dealing with permutattions was basic to the experiment, a sentence of five signa with one hundred and twenty permutations seemed the longest one we could handle easily. The choice was:

You me downtown movie fun?

of which I did not consider the question-aspect, conveyed by the mimicry, to be a signum of the same order: it was neglected insofar as it did not have any influence on the permutations. All five signa were spoken in succession; *you, me* and *downtown* were simultaneously signed; *movie* and *fun* were simultaneously fingerspelled. The sign for *you* was the right index finger pointing at the partner, for *me* it was the same pointing at oneself. The sign for *down town* was the twice repeated sign for *house*: putting the tips of the closed fingers, pointing upward, together. A deaf student addressed his girlfriend this way on a sunny afternoon, May 1959; for the record, I can add that the invitation was accepted and that both went out on a date, including a movie in downtown Washington DC.

Before they received this background information, the student-teachers were asked to interpret the written sentence. All classes, to a total of about one hundred and twenty students, agreed that the sentence consisted of two groups: (*you*+*me*) and (*downtown*+*movie*+*fun*) and that its paraphrase into colloquial English had to be something like:

Why don't you and me go downtown and see a movie for fun?
Would you join me for a movie and some fun downtown?
Are you and me going downtown to see a movie and have fun? etc.

Then they were informed about the background and about the experiment. They were asked to judge each permutation on its sequence of signs and the groups that might appear exclusively, and to make abstraction of the situation in which the original occurred, as well as of possible rhythmical patterning. We presupposed that a decision upon signa constituting a group would imply rhythmical expression in an actual performance. If a permutation remained ambiguous to everyone, it was decided to term it 'not occurring' or esoterically ungrammatical.

There was complete unanimity about all forty-eight permutations which had no insert between *you* and *me*, insofar as *you me* or *me you* were judged to be one group; the other signa *downtown*, *movie* and *fun* were judged to be related to *you me* or *me you* in one way or another as will be worked out hereafter.

1. You me downtown movie fun?
2. You me downtown fun movie?
3. You me movie downtown fun?
4. You me movie fun downtown?
5. You me fun downtown movie?
6. You me fun movie downtown?
7. Me you downtown movie fun?
8. Me you downtown fun movie?
9. Me you movie downtown fun?
10. Me you movie fun downtown?
11. Me you fun downtown movie?
12. Me you fun movie downtown?
13. Downtown you me movie fun?
14. Downtown you me fun movie?
15. Downtown me you movie fun?
16. Downtown me you fun movie?
17. Downtown movie you me fun?
18. Downtown movie me you fun?
19. Downtown movie fun you me?
20. Downtown movie fun me you?
21. Downtown fun you me movie?
22. Downtown fun me you movie?
23. Downtown fun movie you me?
24. Downtown fun movie me you?
25. Movie you me downtown fun?
26. Movie you me fun downtown?
27. Movie me you downtown fun?
28. Movie me you fun downtown?
29. Movie downtown you me fun?
30. Movie downtown me you fun?
31. Movie downtown fun you me?
32. Movie downtown fun me you?
33. Movie fun you me downtown?
34. Movie fun me you downtown?
35. Movie fun downtown you me?
36. Movie fun downtown me you?
37. Fun you me downtown movie?
38. Fun you me movie downtown?
39. Fun me you downtown movie?
40. Fun me you movie downtown?
41. Fun downtown you me movie?
42. Fun downtown me you movie?
43. Fun downtown movie you me?
44. Fun downtown movie me you?
45. Fun movie you me downtown?
46. Fun movie me you downtown?
47. Fun movie downtown you me?
48. Fun movie downtown me you?

The permutation *me you – you me* did not make any difference; neither did the permutations *downtown movie – movie downtown, downtown fun – fun downtown, movie fun – fun movie*; they all appeared to function like the other two signa-groups, quoted above. There was also strong agreement that the following permutations of three signa each, and each occurring four times were identical in meaning:

downtown movie fun
downtown fun movie
movie downtown fun
movie fun downtown
fun downtown movie
fun movie downtown,

and that each such sequence constituted a group. *Downtown, movie* and *fun*, each occurring in initial and in final position twice, were interpreted as a member of the following or preceding group *you me – me you.* It was felt that this extension of the two signa-group into a three signa one did not weaken its coherence at all.

There are thirty-six permutations with one insert between *you* and *me* or *me* and *you.* The first insert is *downtown:*

49. You downtown me movie fun?
50. You downtown me fun movie?
51. Me downtown you movie fun?
52. Me downtown you fun movie?
53. Movie you downtown me fun?
54. Movie me downtown you fun?
55. Movie fun you downtown me?
56. Movie fun me downtown you?
57. Fun you downtown me movie?
58. Fun me downtown you movie?
59. Fun movie you downtown me?
60. Fun movie me downtown you?

Everybody agreed that the insert weakened the coherence of the group and therefore made the sentences more ambiguous. Yet, the majority judged the three signa-group to be intact, regardless of its position at the beginning, at the end or in the middle. In the last case, the sentences were considered to consist of three groups, of which the two one signum-groups *movie – fun* or *fun – movie* embraced the three signa-group.

The minority felt that there was a difference between 49–52 and 53–60. In the former, *you* and *me* were felt to belong to two different groups, whereas the latter were felt to be fully ambiguous, and therefore esoterically ungrammatical. The best explanation I can find is that in the first four cases *you* and *me* or *me* and *you* are in first and third position. The preceding deixis refers what follows to the person pointed at. Deixis+concept(s) thus constitute one group; the repetition with the deixis turning around to the other person creates an opposition. The ambiguity of the other eight permutations suggests that this only occurs with the deixis as first signum of the group. The minority vote might have pointed here at a rule of place for the deixis.

I should like to repeat that the differences of opinion resulted in different rhythmic grouping of the sentences.

The following twelve permutations show *fun* as the insert between *you* and *me:*

61. Downtown movie you fun me?
62. Downtown movie me fun you?
63. Movie you fun me downtown?
64. Movie me fun you downtown?
65. Movie downtown you fun me?
66. Movie downtown me fun you?
67. You fun me downtown movie?
68. You fun me movie downtown?
69. Me fun you downtown movie?
70. Me fun you movie downtown?
71. Downtown you fun me movie?
72. Downtown me fun you movie?

The first six permutations, 61–6, were judged quite differently from the following six, 67–72. Although everybody agreed upon the greater ambiguity caused by the insert, *you* and *me* with the insert *fun* were considered one group by the majority for the former, and members of two different groups for the latter set of permutations. The decisive factor in 67–70 seems again the deixis *you* or *me* in initial position; and also in 71 and 72, with a different sequence. There, the sentence can either have two or three groups, depending on whether *downtown* is part of a three signa-group in first position or a one signum-group. In the latter case, the following two groups with the initial deixis divide the fun and the movie downtown between *you* and *me*. Another matter of order that presents itself differently in the two sets of permutations is the position of *movie*. If *movie* precedes *you* and *me* with the insert *fun*, these three signa are felt to constitute one group by the majority; if *movie* follows they do not. This is true for *movie* only, as well as for *movie downtown* and *downtown movie*. A semantic or psycholinguistic argument might explain why this is so. For its users, and in the situation sketched, *movie* is the central concept, the concrete content of which has a stronger emotional impact than the vaguer notions *fun* or *downtown*. As long as *movie* comes first (61–6), the following *fun* serves as a further elaboration of, let us say, what is expected to result from going to a movie downtown. If *fun* precedes *movie*, it has obviously to form one group with the preceding *you* or *me* (66–70) but this first group is unable to incorporate the following deixis too: *movie* attracts this latter *you* or *me* into the second group. The same is true for 71 and 72, but again with the alternative that *downtown* can be a one signum-group. All this has to be weighed against the increasing ambiguity, of course.

Finally, inserting *movie* between *you* and *me* also results in twelve permutations:

73. Downtown fun you movie me?
74. Downtown fun me movie you?
75. Fun downtown you movie me?
76. Fun downtown me movie you?
77. You movie me downtown fun?
78. You movie me fun downtown?
79. Me movie you downtown fun?
80. Me movie you fun downtown?
81. Downtown you movie me fun?
82. Downtown me movie you fun?
83. Fun you movie me downtown?
84. Fun me movie you downtown?

All informants were unanimous that in the first four instances (73–6) *you me* or *me you* with the insert *movie* constituted one group. The opposite was true for the other eight permutations (77–84): everybody agreed that they consisted of two groups separating *you* and *me*. The difference lies again in the fact that *movie* is the strongest concept, and in the position of the deixis in the group: whether it is initial or not. In 73–6 *movie* occurs in penultimate position and the deixes can not introduce a group. In consequence *movie* constitutes a tight group with *you* and *me*. In 77–80 *movie* can not prevent the following *you* or *me* from starting a new group: the initial position suits these deixes well; moreover, this way the second group has a subject. In 81 and 82 the alternative: the penultimate deixis is the third member of the preceding group with *movie* or it introduces the final group with *fun*, is solved in favour of the latter. And it results in a more acceptable, better balanced sentence indeed. Finally, in 83 and 84 the students felt sure that *you* and *me* belonged to different groups, but they could not decide whether *movie* was a member of the first or the second group. This is understandable, as the central concept *movie* takes center position, flanked by two deixes and the two weaker concepts at the ends. The chances are exactly fifty-fifty.

There was complete unanimity on all the other thirty-six permutations, in which two or three signa were inserted between *you* and *me*: in all of them *you* and *me* definitely belong to two different groups. The reader might want to see for himself:

85. You downtown movie me fun?
86. You downtown movie fun me?
87. You downtown fun me movie?
88. You downtown fun movie me?
89. You movie downtown me fun?
90. You movie downtown fun me?
91. You movie fun me downtown?
92. You movie fun downtown me?
93. You fun downtown me movie?
94. You fun downtown movie me?
95. You fun movie me downtown?
96. You fun movie downtown me?
97. Me downtown movie you fun?
98. Me downtown movie fun you?
99. Me downtown fun you movie?
100. Me downtown fun movie you?
101. Me movie downtown you fun?
102. Me movie downtown fun you?
103. Me movie fun you downtown?
104. Me movie fun downtown you?
105. Me fun downtown you movie?
106. Me fun downtown movie you?
107. Me fun movie you downtown?
108. Me fun movie downtown you?

109. Downtown you movie fun me?
110. Downtown you fun movie me?
111. Downtown me movie fun you?
112. Downtown me fun movie you?
113. Movie you downtown fun me?
114. Movie you fun downtown me?
115. Movie me downtown fun you?
116. Movie me fun downtown you?
117. Fun you downtown movie me?
118. Fun you movie downtown me?
119. Fun me downtown movie you?
120. Fun me movie downtown you?

Of all one hundred and twenty permutations some were more ambiguous than others but none was considered to be esoterically ungrammatical by the majority. Even in 83 and 84 the ambiguity would be solved in actual performance. This leads to a bipartition of all permutations, seventy of which have *you* and *me* in one group, and fifty have not. The former convey the basic meaning that *you* and *me* participate in *movie*, *downtown*, *fun*; the latter convey the opposite. The freedom of order within the sequences is evident. The division can be summarized as follows:

(you me)	24	you ... ... me	12
(me you)	24	me ... ... you	12
		... you ... ... me	3
(you downtown me)	6	... me ... ... you	3
(me downtown you)	6	you ... ... me ...	3
		me ... ... you ...	3
(you movie me)	2	you movie (me	4
(me movie you)	2	me movie (you	4
(you fun me)	3	you fun (me	3
(me fun you)	3	me fun (you	3
	—		—
	70		50

The first decision was unanimous if *you* and *me* had no insert; this was the case in forty-eight permutations; so was the second decision, in thirty-six permutations, if more than one signum came between *you* and *me*. Only in case of a one signum-insert, a choice between the first and the second decision had to be made. Then the initial or non-initial position of *you* and *me*, and the position of *movie* were decisive, as has been worked out.

I think all this proves at least that asking a deaf girl for a date offers some unexpected possibilities.

Further Reading

J. Algeo, 'On defining a stratum', *Emory University Quarterly*, pp. 263–94, 1965.

J. Algeo, 'Stratificational grammar', *South Atlantic Bulletin*, vol. 33, no. 2, pp. 1–4, 1968.

W. S. Allen, 'Structure and system in the Abaza verbal complex', *TPhS*, pp. 127–76, 1956.

W. S. Allen, 'Notes on the Rajasthani verb', *Indian Linguistics*, vol. 21, pp. 4–13, 1960.

Ju. D. Apresian, 'O sil'nom i slabom upravlenij' [On strong and weak governance], *Voprosy jazykoznanija*, vol. 13, no. 3, pp. 32–49, 1964.

Ju. D. Apresian, 'Opyt opisanija značenij glagolov po ix sintaktičeskim priznakam' [An attempt to describe the meaning of verbs by their syntactic signs], *Voprosy jazykoznanija*, vol. 14, no. 5, pp. 51–66, 1965.

Ju. D. Apresian, *Eksperimental'noje issledovanije semantiki russkogo glagola* [Experimental investigation of the semantics of the Russian verb], Moscow, 1967.

R. Austerlitz, E. Hamp, F. W. Householder (eds.), *Readings in Linguistics II*, Chicago University Press, 1966.

K. I. Babitskij, 'O sintaktičeskoj sinonimii predloženij v jestestvennyx jazykax', *IŤI*, no. 6, pp. 29–34, 1965.

Ayo Bamgbose, *A Grammar of Yoruba*, Cambridge University Press, West African Language Monograph Series 5, London, 1966.

Y. Bar-Hillel, *Language and Information*, Addison-Wesley, 1964.

C. E. Bazell, 'On the neutralization of syntactic oppositions', *TCLC*, vol. 5, pp. 77–86, 1949; reprinted in R. Austerlitz, E. Hamp and F. W. Householder (eds.), *Readings in Linguistics II*, pp. 208–15.

C. E. Bazell, 'Syntactic relations and linguistic typology', *CFS*, vol. 8, pp. 5–20, 1949.

C. E. Bazell, 'The fundamental syntactic relations', *Časopis Pro Moderní Filologii*, vol. 33, pp. 9–15, 1949.

C. E. Bazell, *Linguistic Form*, Istanbul University Press (Istanbul Universities Yayinlari, no. 574), 1953.

A. Becker, 'Conjoining in a Tagmemic grammar of English', *MSLL*, vol. 20, pp. 109–21, 1967.

O. Behaghel, *Deutsche Syntax*, Heidelberg, Winter, 1923–32.

J. T. Bendor-Samuel, 'A structure-function description of Terena phrases', *CJL*, vol. 8, pp. 59–70, 1963.

D. C. Bennett, 'English prepositions: a stratificational approach', *J. Linguistics*, vol. 4, pp. 153–72, 1968.

E. Benveniste, 'La construction passive du parfait transitif', *Bull. de la Soc. de Linguistique de Paris*, vol. 48, pp. 52–62, 1952.

E. Benveniste, 'Être et avoir dans leurs fonctions linguistiques', *Bull. de la Soc. de Linguistique de Paris*, 55, pp. 113–34, 1960.

E. Benveniste, 'Fondements syntaxiques de la composition nominale', *Bull. de la Soc. de Linguistique de Paris*, 62, pp. 15–31, 1967.

K. Bergsland, 'Aleut dialects of Atka and Attu', *Transactions of the American Philosophical Society*, vol. 49, pt. 3, 1959.

B. Bloch, 'Japanese syntax', *Language*, vol. 22, pp. 200–49, 1946.

B. Bloch, 'Syntactic formulas for Japanese', *Studies in Linguistics*, vol. 5, no. 1, pp. 1–12, 1947.

B. Bloch, 'English verb inflection', *Language*, vol. 23, pp. 399–418, 1947.

L. Bloomfield, *Tagalog Texts with Grammatical Analysis*, University of Illinois, 1917.

L. Bloomfield, *Language*, Holt, Rinehart & Winston, 1933.

L. Bloomfield, *The Menomini Language*, Yale University Press, 1962.

D. L. Bolinger, 'The comparison of inequality in Spanish', *Language*, vol. 16, pp. 28–62, 1950.

D. L. Bolinger, 'Syntactic blends and other matters', *Language*, vol. 37, pp. 366–81, 1961.

D. L. Bolinger, *Forms of English*, Harvard University Press, 1965.

D. L. Bolinger, 'Adjectives in English: attribution and predication', *Lingua*, vol. 18, pp. 1–34, 1967.

D. L. Bolinger, *Aspects of Language*, Harcourt, Brace & World, 1968.

V. Brøndal, *Essais de Linguistique Générale*, Copenhagen, Munksgaard, 1943.

V. Brøndal, *Les parties du discours* (tr. P. Naert), Copenhagen, Munksgaard, 1948.

V. Brøndal, *Théorie des prepositions* (tr. P. Naert), Copenhagen, Munksgaard, 1950.

G. Calboli, 'I modi del verbo Greco e Latino, 1903–1966', *Lustrum*, vol. 11, pp. 173–349, 1966.

C. R. Carlton, 'Word order of noun modifiers in Old English prose', *J. Eng. Germ. Philology*. vol. 62, pp. 778–83, 1963.

W. Chafe, 'Meaning in language', *American Anthropologist*, vol. 67. no. 5, pt 2, pp. 23–36, 1965.

W. Chafe, 'Idiomaticity as an anomaly in the Chomskyan paradigm', *Foundations of Language*, vol. 4, pp. 109–27, 1968.

A. Chanidze, 'Le sujet grammatical de quelques verbes intransitifs en géorgien', *Bull. de la Soc. Linguistique de Paris*, vol. 58, pp. 1–27, 1963.

S. Chatman, 'Immediate constituents and expansion analysis', *Word*, vol. 11, pp. 337–85, 1955.

C. Church and K. Church, 'The Jacaltec noun phase', *Mayan Studies*, vol. 1, pp. 163–4, 1960.

W. A. Cook, *On Tagmemes and Transforms*, Georgetown University Press, 1964.

W. A. Cook, 'The generative power of a Tagmemic grammar', *MSLL*, vol. 20, pp. 27–41, 1967.

W. A. Cook, *Introduction to Tagmemic Analysis*, Holt, Rinehart & Winston, 1969.

I. Dal, *Kurze deutsche Syntax*, Tübingen, Niemeyer, 1952.

F. Daneš, 'A Three-level Approach to Syntax', *TLP*, vol. 1, pp. 225–40, 1964.

F. Daneš, 'Some thoughts on the semantic structure of the sentence', *Lingua*, vol. 21, pp. 55–69, 1968.

J. Dean, 'Some principal grammatical relations of Bilaan', *Oceania Linguistic Monograph* 3, pp. 59–64, 1958.

G. Deeters, 'Der abchasische Sprachbau', *Nachr. von der Ges. der Wiss. zu Göttingen, Phil.-Hist. Kl.*, pp. 289–303, 1931.

A. W. de Groot, *Structurele Syntaxis*, Den Haag, Servire, 1949.

A. W. de Groot, 'Classification of cases and uses of cases', in M. Halle, H. G. Lunt, H. McLean, C. H. van Schooneveld (eds.), *For Roman Jakobson*, Mouton, pp. 187–94, 1956.

A. W. de Groot, 'Classifications of word-groups', *Lingua*, vol. 6, pp. 113–57, 1957.

A. W. de Groot, 'Subject-predicate analysis', *Lingua*, vol. 6, pp. 301–18, 1957.

P. Diderichsen, *Helhed og struktur*, Copenhagen, Gad, 1966.

S. C. Dik, 'Referential identity', *Lingua*, vol. 21, pp. 70–97, 1968.

K. J. Dover, *Greek Word Order*, Cambridge University Press, 1960.

M. Duff, 'Amuesha (Arawak) syntax I: simple sentence types', *Publicações do Museo Nacional*, Série Linguística Especial, no. 1, pp. 172–237, 1959.

B. F. Elson and Velma B. Pickett, *An Introduction to Morphology and Syntax*, Summer Institute of Linguistics, Santa Ana, 1962.

M. B. Emeneau, *Studies in Vietnamese (Annamese) Grammar*, University of California, 1951.

R. Engel and R. E. Longacre, 'Syntactic matrices in Ostuacan Zoque', *Int. J. Amer. Linguistics*, vol. 29, pp. 331–44, 1963.

C. F. Fillmore, 'The case for case', in E. Bach and R. T. Harms (eds.), *Universals in Linguistic Theory*, Holt, Rinehart & Winston, pp. 1–88, 1968.

J. Firbas, 'Notes on the function of the sentence in the act of communication', *SPFFBU*, vol. A10, pp. 134–48, 1962.

J. Firbas, 'Non-thematic subjects in contemporary English', *Travaux Linguistiques de Prague*, vol. 2, pp. 239–56, 1966.

J. Firbas, 'On the prosodic features of the modern English finite verb as a means of functional sentence perspective', *Brno Studies in English*, vol. 7, pp. 11–48, 1966.

J. R. Firth, 'A synopsis of linguistic theory', *Studies in Linguistic Analysis*, Blackwell, 1957.

J. R. Firth, *Tongues of Men* and *Speech*, Oxford University Press, 1964 (Language and Language Learning, No. 2).

W. Fraser, J. Squair and C. Parker, *Modern Complete French Grammar*, Heath, 1948.

H. Frei, *La grammaire des fautes*, Paris, P. Genther, 1929.

H. Frei, 'The ergative construction in Chinese', *Gengo Kenkyo*, vol. 31, pp. 22–50, 1956; vol. 32, pp. 38–115, 1957.

C. C. Fries, 'On the development of the structural use of word-order in modern English', *Language*, vol. 16, pp. 199–208, 1940.

R. A. Gardner and B. T. Gardner, 'Teaching sign language to a chimpanzee', *Science*, vol. 165, pp. 664–72, 15 August, 1969.

P. L. Garvin, 'L'obviative en Kutenai', *BSL*, vol. 47, pp. 166–212, 1951.

P. L. Garvin, 'Delimitation of syntactic units', *Language*, vol. 30, pp. 345–8, 1954.

P. L. Garvin, *On Linguistic Method*, Janua Linguarum, Series minor, Mouton, 1964.

P. L. Garvin, 'The place of heuristics in the fulcrum approach to machine translation', *Lingua*, vol. 21, pp. 162–83, 1968.

P. L. Garvin, (ed.) *Method and Theory in Linguistics*, Mouton, 1970.

P. L. Garvin and Mathiot, M., 'Fused units in prosodic analysis', *Word*, vol. 14, pp. 178–86, 1958.

A. N. Genko, *Abazinskij jazyk*, Moscow, Isd-vo Akademija Nauk USSR, 1955.

B. L. Gildersleeve (ed.), *The Apologies of Justin Martyr*, Harper & Bros., 1877.

B. L. Gildersleeve, *The Syntax of Classical Greek*, American Book Company, 1911.

B. L. Gildersleeve and G. Lodge, *A Latin Grammar*, MacMillan, 1960.

S. Gili (y) Gaya, *Curso Superior de Sintaxis Española*, Mexico, Minerva, 1943.

H. A. Gleason, *Linguistics and English Grammar*, Holt, Rinehart, & Winston, 1965.

H. Glinz, *Der Deutsche Satz*, Düsseldorf, Schwann, 1957.

S. C. Gudschinsky, 'Mazatec kernel constructions and transformations', *IJAL*, vol. 25, pp. 81–9, 1959.

F. Haase, *Vorlesungen über lateinische Sprachwissenschaft*, Leipzig, Simmel, 1874.

R. A. Hall, 'Roumanian definite article and noun-phrase', *Wortbildung, Syntax und Morphologie: Festschrift für Hans Marchand*, Mouton, pp. 109–14, 1969.

R. A. Hall, 'Comparative reconstruction in romance syntax', *Acta Linguistica*, vol. 11, pp. 81–8, 1968.

M. A. K. Halliday, 'Grammatical categories in modern Chinese', *TPhS*, pp. 178–224, 1956.

M. A. K. Halliday, *The Language of the Chinese Secret History of the Mongols*, Publications of the Philosophical Society, Blackwell, 1959.

M. A. K. Halliday, 'Categories of a theory of grammar', *Word*, vol. 17, pp. 241–92, 1961.

M. A. K. Halliday, 'Class in relation to the axes of chain and choice in language', *Linguistics*, vol. 2, pp. 5–15, 1963.

M. A. K. Halliday, 'Syntax and the consumer', *MSLL*, vol. 17, pp. 11–24, 1946.

M. A. K. Halliday, 'Systemic grammar', *Atti del I e del II Convegno de Studi*, Societá di Linguistica, pp. 23–39, 1969.

M. A. K. Halliday, 'Notes on transitivity and theme in English, Part I, II', *JL*, vol. 3, pp. 37–81, pp. 199–244, 1967.

M. A. K. Halliday and A. McIntosh, *Patterns of Language*, Longmans, 1967.

Z. S. Harris, 'From morpheme to utterance', *Language*, vol. 22, pp. 161–84, 1946.

Z. S. Harris, *Methods in Structural Linguistics*, University of Chicago Press, 1951.

Z. S. Harris, 'Discourse analysis', *Language*, vol. 28, pp. 18–23, 1952.

Z. S. Harris, *String Analysis of Sentence Structure*, Mouton, 1962.

Z. S. Harris, *Discourse Analysis Reprints*, Mouton, 1963.

Z. S. Harris, 'Transformational theory', *Language*, vol. 41, pp. 363–410, 1965.

Z. S. Harris, *Mathematical Structures of Language*, Inter-Science Publishers, 1968.

Z. S. Harris, *Papers in Structural and Transformational Linguistics*, Dordrecht, D. Reidel, 1970.

H. L. Hart, 'Hierarchical structuring of Amuzgo Grammar', *Int. J. Amer. Linguistics*, vol. 23, pp. 141–64, 1957.

F. W. Harwood, 'Axiomatic syntax', *Language*, vol. 31, pp. 409–13. 1955.

A. G. Hatcher, 'The use of the progressive', *Language*, vol. 27, pp. 254–80, 1951.

A. G. Hatcher, *Theme and Underlying Question, Word*, vol. 12 supplement, (Monograph no. 3), 1956.

E. Haugen, 'Directions in modern linguistics', *Language*, vol. 27, pp. 211–22, 1951.

E. Haugen, *The Norwegian Language in America*, University of Pennsylvania Press, 1953.

R. E. Hawkins, 'Waiwai Translation', *Bible Translator*, vol. 13, pp. 164–71, 1962.

A. Healey, 'Notes on Yogad', *Oceania Linguistic Monograph 3*, pp. 77–82, 1958.

P. M. Healey, *An Agta Grammar*, Manila, 1960.

P. M. Healey, 'Levels and chaining in Teleefool sentences', *Linguistic Circle of Canberra Publications*, series B, monograph 5, The Australian National University, Canberra, 1966.

Rev. J. Hinz, *Grammar and Vocabulary of the Eskimo Language* (the Moravian Church, Pennsylvania), 1944.

L. Hjemslev, *La catégorie des cas*, Copenhagen, Aarhus University Press, 1935.

L. Hjemslev, *Omkring Sprogteoriens Grundlæggelse*, Copenhagen, B. Lunos, 1943.

L. Hjemslev, 'Rôle structural de l'ordre des mots', *Journal de Psychologie Normale et Pathologique*, vol. 43, pp. 54–8, 1950.

L. Hjemslev, *Essais linguistiques, TCLC*, vol. 12, 1959.

L. Hjemslev, *Prolegomena to a Theory of Language* (tr. Francis J. Whitfield), University of Wisconsin Press, 1961.

L. Hjemslev, *Language: an Introduction* (tr. Francis J. Whitfield), University of Wisconsin Press, 1970.

C. F. Hockett, 'Two models of grammatical description', *Word*, vol. 10, pp. 210–34, 1954.

C. F. Hockett, *A Course in Modern Linguistics*, Macmillan Co., 1958.

C. F. Hockett, 'Grammar for the Hearer', *Proceedings of Symposia in Applied Mathematics*, vol. 12, pp. 220–36, 1961.

C. F. Hockett, 'Linguistic elements and their relations', *Language*, vol. 37, pp. 29–53, 1961.

C. F. Hockett, *The State of the Art*, Mouton, 1968.

H. Hoijer, see C. F. Voegelin *et al.*

R. D. Huddleston, 'Rank and depth', *Language*, vol. 41, pp. 574–86, 1965.

R. A. Hudson, 'Constituency in a systemic description of the English clause', *Lingua*, vol. 18, pp. 225–50, 1967.

L. N. Iordanskaja, 'O nekotoryx svojstvax previl'noj sintaktičeskoj structury', *Voprosy jazykoznanija*, 4, pp. 102–12, 1963.

A. V. Isacenko, 'Gramatičnost a význam', *A VC, Slavica Pragensia*, vol. 4, pp. 47–51, 1962.

M. Ivić, 'The grammatical category of non-omissible determiners', *Lingua*, vol. 11, pp. 199–204, 1960.

M. Ivić, 'On the structural characteristics of the Serbo-Croatian case system', *International Journal of Slavic Linguistics and Poetics*, vol. 4, pp. 38–47, 1961.

R. Jakobson, 'Beitrag zur allgemeinen Kasuslehre', *TCLP*, 6, pp. 240–88, 1936; reprinted in R. Austerlitz, E. Hamp and F. W. Householder (eds.), *Readings in Linguistics II*, pp. 51–89.

R. Jakobson, 'Shifters, verbal categories and the Russian verb,' Harvard University Russian Language Project, 1957; reprinted in R. Jakobson, *Selected Writings*, vol. 2, Mouton, 1971.

R. Jakobson, *Selected Writings*, vol. 2, *Word and Language*, Mouton, 1971.

O. Jespersen, *Linguistica*, Copenhagen, Levin and Munksgaard, 1933.

O. Jespersen, *A Modern English Grammar on Historical Principles*, 7 vols., Copenhagen, Munksgaard, 1940–1949.

O. Jespersen, *Essentials of English Grammar*, University of Alabama Press, 1964.

O. Jespersen, *Selected Writings*, Tokyo, Senjo Publishing Co., 1964.

O. Jespersen, *The Philosophy of Grammar*, Norton, 1965.

O. Jespersen, *Analytic Syntax*, Holt, Rinehart & Winston, 1969.

W. Jochelson, 'Zametki o fonetičeskix i struckturnyx osnovax aleutskago jazyka', *Izvestija Rossiyskoy Akademij Nauk*, pp. 1031–46, 1912.

M. Joos, 'Description of language design', *Journal of Acoustical Society of America*, vol. 22, pp. 701–8, 1950.

M. Joos (ed.), *Readings in Linguistics I*, Chicago, 1967.

E. H. Jorden, *The Syntax of Modern Colloquial Japanese*, Language Dissertation no. 52, 1955.

H. B. Kerr, 'The case-marking and classifying functions of Cotabato Manobo', *Oceanic Linguistics*, vol. 4, pp. 15–47, 1965.

A. A. Kholodovič, see Xolodovič.

L. D. Kindberg, 'Independent and dependent sentence types in Campa (Arawak)', *A William Cameron Townsend en el Vigésimo-quinto aniversario del Instituto Linguistico de Verano*, Mexico, Summer Institute of Linguistics, pp. 505–18, 1961.

H. V. King, 'Action and aspect in English verb expressions', *English Language Institute Research Papers*, no. 2, University of Michigan Press, 1968.

P. Kiparsky, 'Tense and mood in Indo-European syntax', *Foundations of Language*, vol. 4, pp. 30–57, 1968.

S. Kleinschmidt, *Grammatik der grönländischen Sprache*, Berlin, G. Reimer, 1851.

J. Kurylowicz, 'Dérivation lexicale et dérivation syntaxique', *Bull. de la Soc. de Linguistique de Paris*, vol. 37, pp. 79–92, 1936

J. Kurylowicz, 'Les structures fondamentales de la langue: groupe et proposition', *Studia philosophica*, vol. 3, pp. 203–16, 1939.

S. M. Lamb, 'On alternation, transformation, realization and stratification', *MSLL*, vol. 17, pp. 105–122, 1964.

C. Lancelot, *Nouvelle methode pour apprendre facilement . . . la langue latine*, Paris, 1644.

H. Lander, 'Syntactic patterns in Navaho and Huichol', *Int. J. Amer. Linguistics*, vol. 33, pp. 121–7, 1967.

H. W. Law, 'Morphological structure of Isthmus Nahuatl', *Int. J. Amer. Linguistics*, vol. 24, pp. 108–29, 1958.

H. W. Law, 'Obligatory constructions of Isthmus Nahuatl grammar', *Janua Linguarum*, series practica 29, Mouton, 1966.

H. W. Law, 'The use of function-set in English adverbial classification', *MSLL*, vol. 20, pp. 93–102, 1967.

R. E. Longacre, 'String constituent analysis', *Language*, vol. 36, pp. 63–88, 1960.

R. E. Longacre, *Grammar Discovery Procedures: a Field Manual*, Mouton, 1964.

R. E. Longacre, 'Prologomena to lexical structures', *Linguistics*, vol. 5, pp. 5–24, 1964.

R. E. Longacre, 'Some fundamental insights of Tagmemics', *Language*, vol. 41, pp. 65–76, 1965.

R. E. Longacre, 'Trique clause and sentence: a study in contrast, variation, and distribution', *IJAL*, vol. 32, pp. 242–52, 1966.

R. E. Longacre, 'Reply to Postal's review of *Grammar Discovery Procedures*.' *Int. J. Amer. Linguistics*, vol. 33, pp. 323–8, 1967.

R. E. Longacre, *Discourse, Paragraph, and Sentence Structure in Selected Philippine Languages*, Santa Ana, Summer Institute of Linguistics, 1968.

P. A. Luelsdorff, 'Applicational generative grammar', *Lingua*, vol. 16, pp. 225–37, 1966.

H. G. Lunt (ed.), *Proceedings of the Ninth International Congress of Linguists, Cambridge, Massachusetts, 1962*, Mouton, 1964.

J. Lyons, 'Towards a "notional" theory of the "parts of speech" ', *Journal of Linguistics*, vol. 2, pp. 209–36, 1966.

H. McKaughan, *The Inflection and Syntax of Maranáo Verbs, Publications of the Institute of National Language*, Manila, 1958.

H. McKaughan, 'Overt relation markers in Maranáo', *Language*, vol. 38, pp. 47–51, 1962.

H. McKaughan, 'Sequence of clauses in Tairora', *Oceanic Linguistics*, vol. 5, pp. 1–12, 1966.

K. MacKay, 'The use of the Ancient Greek Perfect down to the second century A. D.', *Bulletin of the Institute of Classical Studies*, 12, pp. 1–21, 1965.

A. Martinet, 'Le genre féminin en Indo-Européen', *Bull. de la Soc. de Linguistique de Paris*, 52, pp. 83–95, 1956.

A. Martinet, 'Reflexions sur la phrase', *Language and Society*, Copenhagen, Berlingske Bogtrykkeri, pp. 113–18, 1961.

V. Mathesius, 'Zur Satzperspektive im modernen Englisch', *Archiv für das Studium der modernen Sprachen und Literaturen*, vol. 84, pp. 200–10, 1929.

V. Mathesius, 'On linguistic characterology with illustration from modern English', *A Prague School Reader in Linguistics*, pp. 59–67, 1963.

D. M. Matson, 'Tagmemic description of agreement', *MSLL*, vol. 20, pp. 103–8, 1967.

A. Meillet, *Linguistique historique et linguistique générale*, Paris, Klinksieck, 1926–1936.

G. A. Menovshchikov, *Eskimosskij jazyk*, Leningrad, 1960.

W. R. Merrifield, 'The Kiowa verb prefix', *Int. J. Amer. Linguistics*, vol. 25, pp. 168–76, 1959.

W. R. Merrifield and B. J. Stoudt, 'Molinos Mixtec clause structure', *Linguistics*, vol. 32, pp. 58–78, 1967.

G. A. Miller, *Language and Communication*, McGraw-Hill, 1951.

A. Mirowicz, *O grupach syntaktycznych a przydawka*, Torun, Nakl. Tow. Naukowego w Torumu, 1949.

T. F. Mitchell, *An Introduction to Egyptian Colloquial Arabic*, Oxford University Press, 1956.

T. F. Mitchell, 'Syntagmatic relations in linguistic analysis', *TPhS*, pp. 101–18, 1958.

R. H. Morse, 'Syntactic frames for the Rawang (Ruwàng) verb', *Lingua*, vol. 15, pp. 338–69, 1956.

C. S. Mundy, 'Turkish syntax as a system of qualification', *BSOAS*, vol. 18, pp. 534–44, 1956.

A. G. Noreen, *Einführung in die wissenschaftliche Betrachtung der Sprache* (ubers. von A. W. Pollak), Halle, M. Niemeyer, 1923.

D. Olmsted, 'Substitution and frequency estimation in linguistic analysis', *Int. J. Amer. Linguistics*, vol. 27, pp. 308–19, 1961.

E. V. Padučeva, 'O strukture mnogosostavnyx složnyx predloženkij v russkom jazyke' [On the structure of compound complex sentences in the Russian language], in *Lingvitičeskije issledovanija po obščej i slavjanskoj tipologij* (Linguistic investigation of general and Slavic typology), T. M. Nikolajeva (ed.), Moscow, 1966.

A. M. Peškovskij, various papers in *Izbrannyje trudy* [Selected Works] Moscow, 1959; also his 'Russkij sintaksis v naučnom osveščenij', [Russian syntax in scientific light (light of science)].

F. R. Palmer, 'Grammatical categories and their exponents', *Proceedings of the Ninth International Congress of Linguistics*, pp. 338–45, 1962.

F. C. C. Peng, 'Structuring of discourse: a neglected aspect of linguistic analysis', *Proceedings of the Eighth Congress of Anthropological and Ethnological Sciences*, pp. 391–4, 1968.

V. B. Pickett, 'The grammatical hierarchy of Isthmus Zapotec', Language Dissertation, no. 56, supplement to *Language*, vol. 36, 1960.

K. L. Pike, 'A problem in morphology-syntax division', *Acta Linguistica*, vol. 5, pp. 125–38, 1945.

K. L. Pike, 'On Tagmemes nee Gramemes', *Int. J. Amer. Linguistics*, vol. 24, pp. 273–8, 1958.

K. L. Pike, 'Language as particle, wave and field', *Texas Quarterly*, vol. 2, pp. 37–54, 1959.

K. L. Pike, 'Dimensions of grammatical constructions', *Language*, vol. 38, pp. 221–44, 1962.

K. L. Pike, 'Non-linear order and anti-redundancy in German morphological matrices', *Zeitschrift für Mundartforschung*, vol. 32, pp. 193–221, 1965.

K. L. Pike, 'Guide to publications related to Tagmemic theory', *Current Trends in Linguistics*, vol. 3, pp. 365–94, 1966.

K. L. Pike, *Tagmemic and Matrix Linguistics Applied to Selected African Languages*, Ann Arbor, University of Michigan Center for Research on Language and Language Behavior, 1966.

K. L. Pike, *Language in Relation to a Unified Theory of the Structure of Human Behavior*, Mouton, 1967.

K. L. Pike, 'Tagmemic theory: an annotated bibliography', *Journal of English Linguistics*, vol. 4, pp. 7–45, 1970.

H. Pilch, 'Comparative constructions in English', *Language*, vol. 41, pp. 37–58, 1965.

R. S. Pittman, 'Nuclear structures in linguistics', *Language*, vol. 24, pp. 286–92, 1948.

A. A. Polebnja, *Iz zapisok po russkoj grammatike*, I–II, Xar'kov, Moscow, Gos. vcebno-pedogog. isd-vo, 1958–1968.

I. Poldauf, 'The so-called medio-passive in English', *PSE*, 13, pp. 15–34, 1970.

P. Postal, *Constituent Structure: A Study of Contemporary Models of Syntactic Description*, Indiana University Press, 1964.

H. Poutsma, *A Grammar of Late Modern English*, Gröningen, P. Nourdhoff, 1928.

D. Premack, 'The education of S*A*R*A*H', *Psychology Today*, vol. 4, pp. 54–8, September 1970.

D. Premack, 'Language in chimpanzee?', *Science*, 172, pp. 808–22, 21 May 1971.

Priscianus, *Institutio*, in M. Hertz and H. Keil (eds.), *Grammatica Latini*, Leipzig, 1859.

R. Quirk, 'Colloquial English and communication', *Studies in Communication*, Secker & Warburg, 1955.

R. Quirk, 'Substitutions and syntactic research', *Archivum Linguisticum*, vol. 10, pp. 37–42, 1958.

R. Quirk 'Towards a description of English usage', *TPhS*, pp. 40–63, 1960.

R. Quirk, *The Use of English*, Longman, 1962.

R. Quirk, 'Descriptive statement and serial relation', *Language*, vol. 41, pp. 205–17, 1965.

P. A. Reich, 'The English auxiliaries: a relational network description', *Linguistic Automation Project*, Yale University, 1968.

P. A. Reich, 'The relational network simulator', *Linguistic Automation Project*, Yale University, 1968.

R. H. Robins, 'Noun and verb in universal grammar', *Language*, vol. 28, pp. 289–98, 1952.

R. H. Robins, *The Yurok Language*, University of California Press, 1958.

R. H. Robins, 'Syntactic analysis', *Archivum Linguisticum*, vol. 13, pp. 78–89, 1961; reprinted in R. Austerlitz, E. Hamp and F. W. Householder (eds.) *Readings in Linguistics II*, pp. 386–95.

Franciscus Sanctius (de Sanctis; Brocensis; Francesco Sanchez de las Brozas), *Minerva, seu de causis linguae latinae commentarius*, Salamanca, 1587.

K. Sandfeld, *Linguistique Balkanique*, Paris, E. Champion, 1930.

S. K. Šaumjan, see S. K. Shaumijan.

J. C. Scaliger, *De causis latinae linguae*, Leyden, S. Gryphius, 1540.

K. H. Schmidt, 'Dativ und Instrumental im Plural', *Glotta*, vol. 41, pp. 1–18, 1963.

E. Schwyzer and A. Debrunner, *Griechische Grammatik*, vol. 2, *Syntax*, Beck, 1950.

C. Scioppius, *Grammatica Philosophica*, Milan, 1628.

H. Seiler, 'On the syntactic role of word order and prosodic features', *Word*, vol. 18, pp. 121–31, 1962.

S. K. Shaumijan, *Applikativnaja porozdnivshachaia model, etc.*, Moscow, Akademia Nauk USSR, 1963.

S. K. Shaumijan, *Transformatsionnyi metod v strukturnoi lingvistike*, Moscow, Akademia Nauk, USSR, 1964.

O. Shell, 'Cashibo II: grammemic analysis of transitive and intransitive verb patterns', *Int. J. Amer. Linguistics*, vol. 23, pp. 179–218, 1957.

N. Ju. Shvedova, see N. Ju. Švedova.

V. N. Sidorov, and I. S. Il'inskaja, 'K voprusu o vyraženij sub'jekta i ob'jekta dejstrija v sovremennom russkom jazyke' [On the question of the expression of the subject and the object of the action in contemporary Russian], Izvestija AN USSR, OL and Ja, Vol. VIII, issue 4, 1949.

H. F. Simon, 'Two substantival complexes in standard Chinese', *BSOAS*, vol. 15, pp. 327–55, 1953.

H. F. Simon, 'Some remarks on the structure of the verb complex in standard Chinese', *BSOAS*. vol. 21, pp. 553–77, 1959.

H. W. Smyth, *Greek Grammar*, American Book Company, 1920.

A. Sommerfelt, 'Sur la notion du sujet en georgien', *Mélanges van Ginneken*, Paris, C. Klincksieck, pp. 183–5, 1937.

L. Spitzer, 'The actor-infinitive construction in Russian and other Indo-European languages', *Word*, vol. 10, pp. 442–56, 1954.

B. M. H. Strang, *Modern English Studies*, St Martin's Press, 1962.

J. C. Street, 'Methodology in immediate constituent analysis', in I. Rauch and C. T. Scott (eds.), *Approaches in Linguistic Methodology*, University of Wisconsin Press, pp. 89–114, 1967.

N. Ju. Švedova, 'Determinirujuščij ob'jekt i determinirujuščeje obstojajel'stro kak samostojatel'nyje rasprostraniteli predloženija [The determining object and the determining adverbial as independent expanders of the sentence], *V. Ja.*, vol. 13, pp. 77–93, 1964.

L. Tesnière, *Esquisse d'une syntaxe structurale*, Paris, C. Klincksieck, 1953.

L. Tesnière, *Éléments de syntaxe structurale*, Paris, C. Klincksieck, 1959.

W. Thalbitzer, 'The absolute and the relative in Eskimo', in (ed.), *A Grammatical Miscellany Offered to Otto Jespersen on his Seventieth Birthday*, pp. 319–29, 1930.

W. Thalbitzer, 'Eskimo', *Handbook of American Indian Languages, Bureau of American Ethnology Bulletin*, vol. 40, pp. 967–1069, 1930.

A. Tobler, *Vermischte Beiträge zur französischen Grammatik*, Leipzig, S. Hirzel, 1902–12.

K. Togeby, 'Structure immanente de la langue française', *TCLC*, vol. 6, pp. 89–142, 1951, *BSL*, vol. 47, pp. 125–7, 1951.

N. Trubetzkoy, 'Le rapport entre le déterminé, le déterminant et le défini', *Mélanges de linguistique offerts à Charles Bally*, pp. 75–82, Geneva, Georg et cie., 1939.

N. Trubetzkoy, R. Jakobson and S. Karcevskij, 'Propositions', *Actes du Premier Congrès International de Linguistiques*, pp. 33–6, 1930.

E. M. Uhlenbeck, *Taalwetenschap*, The Hague, H. K. Smits, 1968.

E. M. Uhlenbeck, 'Some preliminary remarks on Javanese syntax', *Lingua*, vol. 15, pp. 53–70, 1965.

B. E. Ulvestad, 'An approach to describing usage of language variants', *IUPAL* memoir 12 (supplement to *Int. J. Amer. Linguistics*, vol. 22, no. 1), pp. 37–39, 1956.

D. S. Uort, 'Ob otobraženij linejnyx otiošenij v poroždajuščix modeljax jazyka', *Voprosy Jazykoznanija*, no. 5, pp. 46–58, 1964.

J. Vachek, *The Linguistic School of Prague*, Indiana University Press, 1966.

J. Vachek (ed.), *A Prague School Reader in Linguistics*, Indiana University Press, 1964.

J. Vendryes, 'Sur l'emploi de l'auxiliaire "avoir" pour marquer le passé', *Mélanges van Ginneken*, pp. 85–92, 1937.

C. F. Voegelin *et al.*, *Linguistic Structures of Native America*, New York, Viking Fund, 1946.

H. Vogt, 'Un aspect du problème actif-passif dans le verbe', *Journal de Psychologie*, vol. 43, pp. 130–38, 1950.

J. Wackernagel, *Vorlesungen uber Syntax*, vols. 1 and 2, Basel, E. Birkhauser et cie., 1926–1928.

J. Wackernagel, *Kleine Schriften*, Göttingen, Vandenhoeck u. Ruprecht, 1955.

V. Waterhouse, 'The grammatical structure of Oaxaco Chontal', publication no. 19, Indiana University Research Center in Anthropology, Folklore, and Linguistics, (Part II of *Int. J. Amer. Linguistics*) 1962.

A. Weil, *The Order of Words in the Ancient Languages Compared with that of the Modern Languages*, Ginn, 1887.

R. Wells, 'Immediate constituents', *Language*, vol. 23, pp. 81–118, 1947.

A. Wheeler, 'Grammatical structure in Siona discourse', *Lingua*, vol. 19, pp. 60–77, 1967.

W. H. Whitely, 'Some problems of the syntax of sentences in a Bantu language of East Africa', *Lingua*, vol. 9, pp. 148–74, 1960.

E. C. Woodcock, *A New Latin Syntax*, Harvard University Press, 1959.

A. A. Xolodovič, 'Opyt teorij podklassor slov' [An attempt towards a theory of subclasses of words], *V. Ja.*, vol. 9, pp. 32–43, 1960.

V. H. Yngve, 'A model and an hypothesis for language structures', *Proceedings of the American Philosophical Society*, vol. 104, pp. 444–66, 1960.

Acknowledgements

Permission to reproduce the Readings in this volume is acknowledged to the following sources:

1 Munksgaard International Publishers Ltd
2 Modern Language Association of America
3 *Word*
4 Research Centre for the Language Sciences
5 Linguistic Society of America
6 Linguistic Society of America
7 Linguistic Society of America
8 The American Mathematical Society
9 Mouton Publishers
10 Mouton Publishers
11 Georgetown University Press
12 Linguistic Society of America
13 *Word*
14 Linguistic Society of America
15 Pacific Linguistics and Dr Phyllis M. Healey
16 Linguistic Society of America
17 North-Holland Publishing Co.
18 M. A. K. Halliday and Brno University, Czechoslovakia
19 Linguistic Society of America
20 School of Oriental and African Studies and P. J. Honey
21 Mouton Publishers
22 Mouton Publishers
23 North-Holland Publishing Co.

Abrams, N. D., 198, 210–13
Allen, W. S., 19, 84, 88, 89
Anderson, J., 11
Apollonius Dyscolus, 8–11, 14–16, 18
Aristotle, 11, 13, 15
Austerlitz, R., 16
Austin, J. L., 14, 18

Bar-Hillel, Y., 225, 239, 261, 271
Bartholomew, D., 70
Bazell, C. E., 145, 146, 147
Benveniste, E., 82–3, 89
Berge, C., 231, 239
Bergin, O., 131, 133
Bergsland, K., 18, 307
Bever, T. G., 258, 271
Bloch, B., 17, 18, 143, 147
Bloomfield, L., 16, 66, 70, 80
Boas, F., 103
Bolinger, D. L., 17
Bonfante, G., 126, 133
Boole, G., 29
Borko, H., 240
Bossuet, J. B., 55
Bourquin, T., 84, 89
Bradley, F. H., 27
Brandenstein, W., 24, 30
Brockway, E., 80
Bühler, K., 248
Burton, B., 79
Bull, W. E., 34
Bush, R. R., 239

Chanidzé, A., 87, 89
Cherry, C., 240
Chomsky, N., 7, 12, 145–7, 149, 151, 181, 183, 185, 223, 224, 226, 233, 235, 239, 258–60, 265, 269–71
Chrysippus, 8
Clark, S. W., 12, 13, 18
Cocke, J., 230
Cook, W. A., S.J., 16
Cooper, J. F., 69
Couturat, L., 29–30
Curme, G. O., 24, 30

d'Andrade, R., 240
Danehy, J. J., 69
Daneš, F., 248–9, 257
Dean, J. C., 195–6, 198, 200, 205–6, 210–14
Debrunner, A., 83, 90
Deeters, G., 84–7, 89
Delaplane, S., 69
Delbrück, B., 125, 127, 130, 133
Dionysius Thrax, 11, 13, 16, 18
Dirr, A., 85, 89
Dover, K. J., 129, 130, 133
Duff, M., 79, 80
Duijnhouwer, D. J., 243, 245, 247
Dumézil, G., 85, 89
Dyscolus, *see* Apollonius

Eastman, B., 80
Edmundson, H. P., 240
Emeneau, M. B., 276, 286
Engel, R., 79, 80
Erickson, B., 70
Eringa, F. S., 243, 247

Fillmore, C. F., 11, 145, 147
Firbas, J., 249, 256, 257
Firth, J. R., 140, 147
Fodor, J. A., 158, 271
Fowler, H. W., 40
Fraser, B., 258
Fries, C. C., 68, 69

Gaifman, H., 225, 230, 233–4, 239
Galanter, E., 239
Gardner, R. A., and Gardner, B. T., 273
Garrett, M., 258, 271
Garvin, P., 17, 238, 239, 288–94, 296, 300
Gaya, Samuel Gili, *see* Gili (y) Gaya, S.
Genko, A. N., 85, 89
Gildersleeve, B. L., 8
Gili (y) Gaya, S., 38, 50
Ginsburg, S., 236, 239
Girard, Abbé, 26, 30

Gonda, J., 125, 126, 133
Gorn, S., 231, 239
Gougenheim, G., 55, 65
Greenberg, J., 82
Greene, H. R., 12–13, 18
Greibach, S. A., 236
Gruber, J., 11

Haas, Mary, *see* Swadesh, M. H.
Hahn, E. A., 16, 18
Halle, M., 269, 271
Halliday, M. A. K., 12, 16–17, 140–1, 144, 147, 149, 248–50, 256–7
Hammerich, L. L., 303, 307
Hamp, E., 16, 226, 239
Harmer, L. C., 55, 65
Harris, Z. S., 7, 17, 18, 103, 143, 147, 151, 185, 223
Hatcher, A. G., 11, 16, 18, 37, 40, 44, 47, 63, 65, 273
Hawkins, R. E., 218, 222
Hays, D., 17, 228, 230, 233, 236, 240
Healey, A., 197, 214
Healey, P., 149, 197, 214
Henrici, A., 248
Hermann, E., 24, 30
Hinz, J., 301, 307
Hjemslev, L., 85, 90, 141, 147, 224, 240, 291, 295, 300
Hockett, C., 17, 18, 45, 50, 69, 141–3, 147, 239
Höffding, H., 25, 30
Hofmann, J. B., 129, 133
Hollenbach, B., 66, 80
Honey, P. G., 18
Huddleston, R. D., 248, 250, 257
Hudson, R. A., 248, 255, 257
Hymes, D. H., 124, 133

Ihm, D., 232, 240
Ingstad, H., 307
Iordanskaja, L. N., 224, 240
Irish, F. V., 12, 18
Ivanov, V. V., 125, 133–4
Ivić, M., 136, 137, 138–9

Jacobsen, W. H., 236, 240
Jakobson, R., 82
Jakovlev, N. G., 85, 90
Jespersen, O., 17, 19, 26, 30, 40, 50, 51, 307
Jochelson, W., 305, 307
Joos, M., 17, 19

Kahane, H., and Kahane, R., 33, 50
Kant, I., 26
Kay, M. J., 239
Kellogg, B., 11–12, 19
Kellogg, S. H., 88, 90
Kindberg, L. D., 78, 80
Kleinschmidt, S., 84, 90, 301–3, 307
Kroll, W., 129, 133
Kurylowicz, J., 125, 126, 132, 134

Lamb, S. M., 149, 235–9, 240
Laroche, E., 128, 134
Le Bidois, R., 55, 65
Lecerf, T., 230, 232, 239, 240
Lees, R. B., 145–7
Leibniz, G. W., 29–30
Lê văn Lý, 276–7, 285–6
Lewis, S., 69
Longacre, R. E., 79–80
Loriot, J., 66, 80
Lounsbury, F., 258, 261
Luce, R. D., 239
Lý, Lê Văn *see* Lê văn Lý
Lyons, J., 16, 82

Maclay, H., 80
March, F. A., 13, 19
Marius Victorinus, 16
Marks, L. E., 259, 272
Martinet, A., 142, 147
Marty, A., 24
Mathesius, V., 248
Mathiot, M., 288, 300
Matthews, P., 16
McCawley, J., 258, 271
McKaughan, H., 16, 197, 203, 205, 212, 214
Mel'cuk, I. A., 16

Menovshchikov, G. A., 307
Miller, G., 115, 123, 224, 226, 239, 259, 272
Mirowicz, A., 136, 139
Monro, D. B., 130
Montano, Severino, 93, 95
Mullikin, T., 228

Nehring, A., 24–5, 30
Newman, S. S., 84, 90
Nida, E. A., 69, 75, 81

Otten, H., 129, 130, 134

Palmer, F., 16
Palmer, H. E., 39, 50
Perles, M., 225, 239
Pickett, V. B., 78–9, 81, 203, 214
Pike, K. L., 17, 18, 140, 144, 147, 198, 202, 214, 273
Pittenger, R. E., 69
Plato, 9, 11
Poldauf, I., 248, 257
Postal, P. M., 146, 147, 226, 258, 260, 261, 272
Potebnja, A. A., 136, 139
Poutsma, H., 31, 35, 39–41, 46, 48
Premack, D., 273
Priscian, 8, 10, 14, 19

Quintilian, 11

Reed, A., 11, 12, 19
Reich, P. A., 149, 258, 262, 272
Reichling, A., 19, 308
Rinehart, M. R., 69
Robins, R. H., 83, 90, 243, 247
Romney, A. K., 240
Rose, G. F., 236, 239
Rosenbaum, P. S., 12, 19
Ross, J., 258

Sanctius (de Sanctis, Sanchez de las Brozas), 15, 19
Sapir, E., 69, 168, 185, 191, 194
Saussure, F. de, 9, 19
Sauvageot, A., 302, 307
Schuchardt, H., 86
Schuler, E. von, 128, 134
Schulz-Lorentzen, C. W., 84, 90
Schwyzer, E., 83, 90, 130, 134
Shamir, E., 225, 239
Sigwart, C. von, 26
Simon, H. F., 281, 286
Sleator, M. D., 80
Slocum, M. C., 79, 81
Slotty, F., 25–6, 27, 30
Smyth, H. W., 9, 10, 19
Sommer, F., 130
Sommerfelt, A., 86, 90
Spencer, H., 41
Spitzer, L., 59
Stephanos, 13, 19
Storrs, R. S., Jr., 13
Strunk, W., Jr., 69
Swadesh, M., 84, 90, 103, 110, 114
Swadesh, M. H., 110, 114

Tervoort, B. T., 273
Tesnière, L., 12, 17, 19, 149
Thalbitzer, W., 84, 90, 301–3, 307
Thorne, J. P., 15, 19
Trager, G., 83, 90, 103, 110, 114
Trubetzkoy, N. S., 85, 90, 136, 139, 140
Tryphon, 15
Tschenkéli, K., 86–7, 90

Ueberweg, F., 26
Uhlenbeck, E. M., 19
Uldall, H. J., 141, 147
Uren, O., 250, 257

Vachek, J., 248–9, 257
Van der Dussen, W. J., 245, 247
Văn Lý, Lê *see* Lê văn Lý
Van Syoc, W. B., 243, 247
Varro, M. Terentius, 83
Vendryes, J., 82, 90
Victorinus, M. *see* Marius Victorinus

Voegelin, C. F., 83, 90, 103
Vogt, H., 82, 86–7, 90, 302, 307

Wackernagel, J., 8, 125–132, 134
Waterhouse, V., 78, 81, 273
Watkins, C., 18
Weil, H., 17
Wells, R., 17
Wharton, E., 69
White, E. B., 69
Whitehall, H. E., 75
Whitfield, F., 295, 300
Whorf, B. L., 110, 114, 194
Wolfenden, E., 197, 212, 214
Wonderly, W. L., 83, 90
Wurm, S. A., 218

Yngve, V., 17, 265, 270, 272

Žirkov, L. I., 85, 90

Abaza, 84–5, 88
Abilitative, 216
Abkhaz, 85
Ablative case, 10, 15, 83, 192
Absent *v.* present, 107
Absolute, absolutive (suffix, case, object) (abs.), 84, 109, 301–3, 305–7
Accent (mark), 93
Accented, accentual, accentuation, 9, 93, 127
Acceptable, acceptability, 259, 260, 270–1, 308, 315
Accepted, 228–30
Accompaniment, 215
Accusative, 9–10, 14, 135, 137–8
Action, 98, 190, 197, 248
Active voice, 7, 83, 114, 122, 144–5, 155–6, 190, 194, 197, 212–13, 243–7, 256, 301
Activity, 56–9, 197–8
  focus (relation), 197–9
Actor (Ac), 82–3, 144, 197–9, 200, 202–6, 211–13, 218, 248–50, 253, 255, 301–3
  focus (af), 197–200, 203–4, 207, 209, 212
  marker, suffix, tagmeme, 202–3, 205, 213, 301–3
Addition, 72–5
Additive sentences, 67
Address, terms of, 102
  see also Vocative
Adequacy, adequate, 235, 259, 271
Adjective (A), 10, 12, 14–15, 38, 40–7, 49, 103–5, 107–9, 112, 120, 136–8, 152–4, 156–7, 161–2, 165, 167–8, 171–2, 175, 180–1, 183, 191–2, 223, 241–2, 256, 286
  clauses, 23–4, 35
  phrase (A-phrase), 156, 162, 164, 168–70, 173
Adjectivization, 112
Adjunct (adjective, participle), 34, 40–1, 44–6, 48, 141, 256
Adverb (D), 10, 34–9, 43, 46, 120, 138, 152, 155–6, 159, 171, 180–1, 225, 242, 297
Adverbial (element, question, adjunct), 48, 52, 56, 60, 155, 252
  clauses, 10, 23, 24, 37–8
  conjunction, 39
  phrases, 38, 166, 241, 243
Affirmations, 15
Affix, affixation, 151–2, 155, 158, 174–5, 218, 231, 237, 245
African languages, 149
Afterthought, 80
Agent, agency, 56, 62, 83, 84, 86, 96–100, 144, 190, 213
Agglutinative, 111
Agreement, agree, 10, 84, 153, 225, 237, 290–93, 301
  see also Concord
Alaskan, Eskimo, 112, 301–5
Aleut, 273, 304–7
Algonkian, 104, 109, 113
Algorithm, 230
Allomorph, allomorphic, 194, 202–4, 210, 290
Allotagma, allotagmatic, 200, 202, 208
Alphabet, 224, 233, 235–7
Alternants, 210, 213, 303–4
Alternative (classes, order), 113–14, 170–1
Ambiguity, ambiguous, 9, 110, 200, 206, 208, 212–13, 216, 220, 225, 230–4, 298, 312–14, 316
American Indian, 83, 103, 106–7, 149, 195
Amplified sentence, 126–7
Amuesha, 79
Analysis, analytic, 14, 54, 151–2, 157, 177, 180, 186–7, 194, 196, 203, 205, 211, 219, 249, 275–6, 287–95, 297, 299, 301, 304, 307–8
Anaphora, anaphoric, 27, 29, 66–7, 70, 95
Anaphoric (An), 250–6
Anatolian languages, 125–6
Answer, 14, 55, 67, 72, 78, 165

Antecedent, 66, 70, 162, 168, 304
Aorist, 14, 86–8
Apophony, 131
Apposition, appositional, 10, 44, 169, 220
Arabic, 111–12
Arbitrary rule, 142
Archaic, archaism, 129, 136
Area, linguistic, 296
Armenian, 82
Arrangement, 288
  see also Order, Sequence
Article (T), 9–10, 15, 68, 104, 109, 112, 117, 152–3, 156, 161, 173, 193–4
Aspect, aspectual, 53, 58, 86–8, 110–14
Assertion, 14, 156, 165–6, 169, 256
Assignment function, 225–32
Associated clause, 216, 218–21
Athapaskan, 124
Atka (Aleut), 305–6
Atonic, 132
  see also Enclitic, Proclitic
Attitude of speaker, 248, 256–7
Attribute, 58, 97–101, 276
Attributive (adjective, expression, genitive), 24, 38, 44, 94, 96–7
Augment, 133
Autonomous, autonomy, 23–4, 27, 190, 192–4
Autosemantic, 24–5
Auxiliary verb (v), 56, 68, 82–3, 88, 112, 145, 152, 158–9, 162, 167, 173–5, 194, 234, 241
  alphabet, 124–5, 233
  character, element, 225, 227–8, 231, 234
Aztec, 67, 80, 108–9, 111, 113–14
Azteco-Tanoan, 110

Baltic, Balto-Slavic, 125, 132
Bantu, 113
Base, 178, 210
Basic (sentence, structure), 241–7
Benefactive, 218, 245
Beneficiary, 190, 215, 220, 249
Bilaan, 195–213
Bilateral dependence, 291
Bound (form, morpheme, unit), 66, 104, 136–7, 160–1, 164, 172, 275–6
Boundary, bounded, 130, 183, 250, 253–4, 270, 284, 288–90
Bracketing, 8, 255
  see also Parenthesization
Branch, branching, 8, 117–20, 169–71
Breadth hypothesis, 270
Burushaski, 125

Caesura, 130
Campa, 67, 78
Carrier (morph), 158–9, 162
Case, 10–11, 23, 82–5, 107–8, 110, 114, 136, 190–92, 194, 294, 301, 303, 306
  category, form, 136–8
  ending, suffix, 88, 109, 294
  grammar, 11
  meanings, 138
Categorical grammar, 17
Categorical propositions, 29
Categoric plural, 96
Category, 11, 16, 51, 64–5, 83, 93, 103–14, 124, 136–7, 140, 145–6, 231, 237, 241, 243–4, 247–8, 277, 283, 291, 295
Caucasian languages, 84–6, 149
Causal clause, cause, 10–11, 248
Causative (form, voice, phrase), 39, 105, 107, 302
Celtic, 125
Central American Indian languages, 149
Central (element, unit, concept), centre, 226–7, 231, 314–15
Central embedding, 259, 264–6, 271
Character (auxiliary, terminal), 224–30
Characterization, characterize, 43–5, 223–31
Chimpanzee, 273
Chinese, 11, 109, 286
Chontal, 78
Circumstance, circumstantial clause, 28, 248

Citation (form, matrix, paradigm), 195, 200–10
Class (of forms, morphemes, words), 104–8, 113, 132, 151–61, 163, 171, 183–4, 199–200, 203–6, 208, 223–8, 242, 245, 247, 275–86, 289–94, 296–8
  marks, meaning, 109–10
  membership, 105, 181, 296–7
  order, sequence, 103, 152
  see also Word-class, Form-class
Classification, 18, 24, 52, 54–5, 59, 61, 151, 276–7, 289–90
Clause, 10, 23–30, 38, 40, 61, 66–7, 69, 79, 101, 117, 121, 126–7, 130, 164, 195–9, 203–13, 215–22, 241, 248–57, 263–4, 271, 284, 294–9, 305–6
  base, 210
  constituents, members, 28, 295, 298
  coordination, 299
  final, terminal, 210, 217
  formula, 203, 205–6
  functions, 294–9
  level, 215–16, 220, 221
  structure, paradigm system, 210, 248–9
  subtype, variant, 205, 216–17, 219
  types, 16, 195, 197, 215–22
Close attribute, 99–100
Close juncture, 94, 99–101
Cluster of morphemes, 289–90
Cognate, 125, 304–6
Cognitive, 248, 255
Coherence, cohesion, cohesive, 251–3, 313
Colligation, 281
Colloquial, 114, 311
Combiners, combining, 179–80, 184
Comma intonation, 166, 169, 175
Command, 15, 35, 256
Comment, 72, 74, 78, 256
Common Gender Class, 107
Common noun, 105
Communication, communicative function, 31, 187–8, 249, 308–9, 311
Commutable, 186
  see also Substitutable
Comparative, comparison (adjectives, clauses), 10–11, 13
Comparative and historical (grammar, linguistics, method), comparison, 18, 82, 124–6, 131, 304, 307
Comparative syntax, 124–33, 307
Competence, 259, 269
Complement, 9, 26, 29, 48, 197–200, 202–6, 208, 210–13, 215, 250, 255, 297–8, 301, 305–6
Complete *v.* incomplete sentences, 68
Completive sentences, 66–70, 72, 75–80
Complex (sentences, structures, suffixes), 10, 77, 164, 181, 303
  information, structure, modality, theme, 251–4
  word, 276
Component of a grammar, 7, 224, 241, 248–9, 257
Component of a linguistic unit, 191, 196, 198–9, 205, 211, 243, 255
Componential analysis of function, 256
Composite (construction, expression, tagmeme), 153, 205, 223
Composition of morphemes, 143
Compound, compounded, compounding, 9, 46, 61, 105, 127–8, 171, 311
  sentence, 10, 113
  tense, 52
  verb, 85, 128, 131–2, 171, 233
  words, nouns, 9, 180, 311
Comprehension of sentences, 270
Computational linguistics, 17
Computer, 17, 116, 258, 267
Concatenate, concatenation, 178, 183, 253, 262–3, 266–7
Concept, 9, 314–15
Concessive clause, 10, 16
Conclusion, 54
Concomitant, 78
Concord, 10, 11, 82, 88, 144, 191–3, 213
  see also Agreement

Condition, conditional (clause, proposition, mood), 10, 28–9, 37, 302, 306
  see also Unfulfilled event
Configurational, configurative, 103–5, 113
Conflate, conflation, 141–2, 250, 253–4
Conjoin, conjoining, 12
Conjugations, 113
Conjunct, 251–4
Conjunct attribute, 98–9
Conjunction (C), 15, 23–4, 26–7, 127, 129, 152, 164, 166–8, 169, 172, 235, 262, 270, 298
Conjunction, construction, concatenation, 135, 137
Conjunctional clause, 23
Conjunctive (Co), 79, 250, 253–4, 302
  adverbs, 37
Connective (function, particle, suffix), 67, 100, 113, 126–33, 169, 191, 220–21
Connector, 224
Connotation, 112
Consequence, 248
  see also Result
Constellation, 291
Constituency, constituent (analysis, structure), 8, 12–13, 24–8, 115–17, 121, 129, 145–6, 160, 183, 204–5, 233–4, 255–6, 260, 291, 295
Construction, construct, 52–4, 61–4, 66, 83, 96–7, 116, 120, 129, 152–63, 170–71, 176, 179, 183, 195–8, 200, 204, 210, 221–2, 232, 238, 242, 258, 261, 265, 270–71, 301–3, 305–6
Constructional, 53–4, 62, 168, 173
  equivalence, 153
  position, 168, 173
  rules, 184
  status, 153–4, 163, 168
  usage, 53–4
Contemporative mode, 302, 304
Content, semantic, 130, 193, 219
Context, contextual, 25, 32, 35, 54, 57, 68, 72, 104, 192–3, 219, 309
Context-bound, 242
Context-free (grammar, language, rules), 223, 233, 235–6, 261
Context-sensitive, 236, 261–2
Contiguous elements, 31
Continuous style, continuity, 54, 132–3
Contour of intonation, 80, 126, 268
Contrast, 33–5, 39–43, 46, 49, 67, 169, 195, 197, 200, 203–4, 244
Contrastive, 34, 47–8, 195–7, 200, 251–2, 254
  stress, 34–5, 39, 49, 166, 169
  variation, 142, 145
Conversation, 35, 67, 75, 215, 249
Co-occurrence, co-occur, 151–6, 158, 160–3, 169, 178–83, 238, 273, 287, 308
Co-occurents, 151, 160, 183
Coordinate, 23, 29, 164, 221
Coordination, 23, 120, 268–71, 299
Coordinator, coordinating conjuction, 100, 169
Copula, 13, 61
Copulative, 107, 129, 222
Core member, 297
Corpus, 69
Correction, 54
Correlation, morphological, 308
Correspondence (of form & meaning, of cognates, of structures), 234, 294, 296, 304
Co-variance, 287
Covert (grammatical categories, classes), 104–9
Cross reference, cross referent (cr), 205–7, 209, 210
Cryptotype, 107–8, 111
Cycles of analysis, 288–93, 297, 299
Cypriote, 130
Czech, 135, 257

Danish, 190
Data, linguistic, data-base, 53, 153, 215, 239, 270, 287, 291–2, 297, 299
Dative, 9, 10, 83, 86, 87, 189, 191–2, 248
Deaf, deaf-mute, 273, 308–13

Decision procedures, 224, 226–7, 228, 230, 236
Declarative, 79, 243–7, 250–4, 256–7
Declension, 113
Deduction, 54
Deep structure, 7, 13, 14
Definite article, definiteness, 6, 63, 68, 83, 104, 117, 245, 257
Definition, 8, 10, 11, 25, 103, 233, 288, 293, 299
Deictic, deixis, 128, 313–15
Delaware, 83
Deletion, transformation, 7, 11, 13, 15, 241
Deliberative questions, 9
Delicacy, 242, 254
Demonstrative, 68, 96, 297–9, 305
Denials, 15
Denotation, 197
Dependence, dependency, 12, 80, 180, 223–39, 270, 290–95, 297, 299
  grammar, theory, 12, 223–39
  rule, structure, 225, 228, 230, 233, 270
  trees, D-trees, 12, 224, 235–8
Dependent, 66–9, 157, 169, 225–7, 229, 233, 238, 295, 297–8, 306
  clauses, 61, 66, 117, 132
  marker, 79
  mode (depend.), 216, 302–4, 306
  sentences, 66–80
  see also Subordinate
Depth of embedding, 118, 122, 263–4, 271
  hypothesis, 115–22, 265, 270
Derivation, derive, 9, 111–12, 243
Derivational, 243, 267, 305
  history, 269
  suffix, 302
Derivative, 13, 98, 306
Description, 54, 74, 116, 144, 151, 206, 208, 241–2, 254, 287, 299
Descriptive, 74, 114, 250–54, 302
  adjective (A), 34, 168
  adverb, 35
  analysis, grammar, linguistics, 24, 51, 183, 304
Desiderative clause, 216–18, 220–21
Destination, 215
Determination, determine, determining, 25–6, 291, 311
Determiner (D), 135–8
Deverbal, 243
Deviant, 308
Diachronic grammar, 138
  see also Historical
Diagnostic environment, 151–2
Diagram, diagramming, 11–13, 230–32
Dialect, 133, 136, 302
Dictionary, 236–7
  see also Lexicon
Digression, 76
Direct case, 88
Direct imperative, 216–17, 219
Direct object, 11, 84–8, 135, 260, 297, 309
Direct quotation, 79
Direct speech, 218
Discontinuity, discontinuous (constituent, morpheme), 121–2, 141, 142, 150, 157, 160, 163, 191, 232
Discourse, 32, 49, 66–7, 80, 215, 242, 249, 255, 257
  theme, 251–2, 254
Discovery procedures, 15, 51, 149, 223
Discussion, 215
Disjunction ('or'), 262, 265
Disjunctive ('but'), 79
Distinctive units, 187
Distribution, 187, 220, 223, 287, 289
  class, 138, 213, 289
Distributional (analysis, technique), distributionalist, 138, 186–7, 287–90, 293–4, 299
Ditransitive verb, 241, 243, 245, 247
Domain, 155–6, 181, 183, 295
Double articulation, 187
Double function, 203
Drop transformation, 176, 182
  see also Deletion
Dropping test, 291–2, 295, 297
Dual (du.), 15, 301–2

Durative, 33, 110
Dutch, 310–11
Dutch School, 17

Eastern Eskimo (EE), 303, 306
Economy, 146, 193, 232, 235
Effective, 252
  mode, 83–4
Element, 32, 52, 104–5, 110, 124, 127, 131, 140–42, 158, 169, 171, 175, 179, 180, 183, 186–8, 195, 224–5, 231–3, 237–8, 241–2, 250, 264–6, 269, 275, 287, 307
Elementary (sentences, transformations, units), 176–9, 184, 289
Elicitation, elicit, 18, 290, 295–6
Ellipsis, 13, 15, 303
  see also Deletion, Drop transformation
Elliptical, 252
Embed, embedding, 146, 250, 255, 258, 264–5
  see also Imbedded
Emic, 196, 198, 205–6, 208, 210, 213
Emphasis, 67, 78, 156, 197–9, 202–4
Emphasized, emphatic (e), 96, 158, 199–201, 207, 209–11, 243
Emphatic pronouns, 203, 206
  stress, 158
  tagmeme complex, 211
Empty (word, preverb), 109, 128, 131
Enantiomorphism, 114
Enclitic, 94, 96–7, 99–102, 124–33, 304
  connective (N), 127, 130, 132
  pronoun, pronominal element (E), 96, 124, 127, 132, 304
Endings, 108, 192, 265
Endocentric groups, 283
English, 7, 10, 11, 15, 23, 31, 34–5, 41–2, 45, 50, 56, 67–70, 72, 77, 82, 103–8, 110, 112–15, 117, 119–22, 144–6, 151–2, 164, 170, 182, 188, 190–91, 194, 197–8, 202–3, 212–13, 225, 233, 248–9, 256–7, 258–61, 265, 276, 285, 308–11
Entry condition, 253–4, 257
Enumerate, 261
Enumerative contrast, 49
Environments, 151–5, 157–60, 172, 177, 183, 185, 257
Episodes, 80
Equational (clause, sentence), 15, 216, 219, 242, 244, 247
Equative, 256–7
Equipotent, 232–8
Equivalent, 232–4
Ergative (case, construction), 85, 302
Eskimo, 84, 112, 273, 301–7
Esoteric, esoterically ungrammatical, 308, 312–13, 316
Ethic dative, 248
Etic, 196, 198, 205–6, 210
Exclamations, exclamatory (sentences, intonations), 10, 68, 78–9, 180, 215, 219
Exclusive (we), 94–5, 296
Existence (sentence-theme), 55
Expansion, 66, 72–5, 116, 170, 173, 205, 210, 241–2
  see also Extension
Expectancy, 140
Experiment, experimental, 256, 270
Explanation, explanatory, 13, 15, 54, 79, 304
Explicit, 13–14
Exponent, 140–1, 143–6
Expressions, 93, 194, 230–31, 293
  see Phrase, Group
Extension, 221
  see also Expansion
Extensive, 250–4
Extraposition, 12

Factitive phrase, 39
Falling terminal contour, 268
Far Eastern languages, 149
Favourite sentence types, 241–2
Features, 216, 295
Feedback signal, 262–3, 265, 267
Feminine gender, 105, 107

Field work, 296
Figurative sense, figure of speech, 10, 40, 46
Filler of a tagmemic slot, 205–8
Final relation, 248
  see also Purpose, Purposive
Final (position, place), 126–9, 132, 250, 284, 313
  fall, intonation, juncture, rise, 80, 243–4
  particle, marker, 79, 284, 286
  verb, predicate, 305–6
Finite verb, 13, 124–5, 130, 132, 137
  element, language, 184, 264
  state (device, diagram, machine), 122, 236, 258, 260–66
  string, 228
  system, 260, 271
First person (1 p.), 14, 86–7, 94, 216, 281, 220, 302–5
  singular (1 p.sg.), 14, 87, 220, 303–4
  plural (1 p.pl.) inclusive, exclusive, 68, 296
First *v.* last, 31
Fixed-order, 126, 289
Flexional, 138
Focus, 197–206, 210–13, 252
  complement (fc), 197–200, 202–6, 208, 210–13
Form, 10, 23, 26, 51, 53, 62, 66, 114, 138, 188–90, 202, 245, 287–90, 296, 301
Formal, 82, 104, 109, 113, 133, 140, 145, 294
  discourse, 68
Formalism, 224
Formatives, 7
Form-meaning (technique, correspondence, cycle), 289, 294, 296, 299
Formula, 124, 153, 155, 189, 203, 205
Fourth person (4p), 301–3, 305–6
Frame, 273, 289–91, 298–9
Free (form, morpheme, unit), 66, 136, 176, 275–6
  position, order, 35, 129, 308–9, 311, 316
  variation, variant, 183, 237
French, 10, 82, 109, 111–12, 189, 194
Full expressions, sentences, 95–7, 102, 306
Full words, 94–6, 101–2, 109
Function, 10, 23, 110, 135, 188–93, 202, 248, 250, 256–7, 289, 291, 293–9, 304
  dependence, 291, 295, 297–8
  words, 308–9
Functional, 103–4, 133, 192–4, 203, 293
  load, 210
  moneme, 190–4
  sentence analysis, functional sentence perspective (FSP), 248–9, 256–7
  slot, 196
Functionalist, 248
Functive, 141
Fused units, 288, 290
Future tense, 16, 100–101, 216, 303, 309

Gender (class, suffix), 15, 105–7, 109, 113, 153
Generate, generation, 184, 223, 228, 267
Generative function, characterization, procedure, 224, 226–8
Generative grammar, 241
Generative semantics, 22
Generic category, 114
Genetic, filiation, 125
Genetivus (possessivus, qualitatis, temporis), 135–6
Geneva school, 17
Genitive, 9–10, 15, 82–3, 85, 136
Gentilic plural, 104
Georgian, 86–9, 302
German, 189
Germanic, 8, 132

Gerund, 38, 46, 137
Gesture, 308
Given, givenness, 56–9, 63–4, 249–50, 255–6
  see also Known
Glossematic, 141, 224
Goal, 98, 144, 197–200, 202–4, 206, 208, 211–13, 249, 255
  -focus (gf), 197–201, 203–7, 211–13
Gothic, 132
Governing (element, member), 9, 28, 224–5, 227–8
Government, 11, 290–93
Governor, 225, 233, 238
✓ Gradation (of position, meaning), 37
Grammar, 7–10, 103, 105, 110, 113–14, 116–18, 120–21, 135–6, 142, 145–6, 155, 157–8, 163, 171, 181, 183–4, 204, 223–38, 247, 248–50, 253–4, 257–9, 262–3, 265, 267, 286, 311
Grammarians, 9–11, 13, 16, 23, 55, 104, 114, 122, 273, 277
Grammatical, 8, 14, 56, 66, 80, 105–11, 114, 115, 122, 135–7, 140–6, 153–7, 184, 191–2, 194, 238, 259, 269, 275, 277, 286, 301, 308
  cases, 84–5
  category, class, 82, 103–16, 136, 146, 276, 291
  elements, units, 135, 142, 171, 191, 237, 307
  features, meanings, 105, 137, 218, 293, 296
  form, 23–4, 103, 110
  level, system, 138, 225, 275, 285
  relation, 140, 154, 168, 184
  subject, object, 56, 82, 305
Grammaticality, grammaticalness, 80, 238, 259–60
Graph, 237
Grapheme, 231
Greek (Gk.), 8, 10–11, 13, 15, 17, 67, 80, 83, 109, 111, 124–6, 129–30, 132, 133
Greenlandic, Eskimo (Grl.), 301–3, 305
Greeting, 79
Group of words, 280–1, 283, 308, 310–16
  see also Phrase, Expression
Gujarati, 88

Half-transitive verbs, 301, 306
Harauti, 88–9
Head, headship, 12, 66, 98–9, 101, 136, 153, 156, 173–4, 211, 292–3
Hearer, 32, 218, 248
  see also Listener, Interlocutor, Second person
Hebrew, 111, 113
Hierarchy of parts of speech, 120
Hieroglyphic, 'Hittite', 128
High (pitch, stress), 93, 99
Hindi, 88
Historical (syntax, reconstruction), 24, 124–5, 159, 307
  see also Comparative
Hittite (Hitt.), 124, 126–33
Hokan, 111
Homeric, 130
Homonymous, homonymy, 89, 173–5, 179, 181, 309
Homophones, homophony, homophonous, 213, 285
Hopi, 105, 107–8, 111–14
Human (subject, object), 58, 62
Hungarian, 83
Hypotaxis, 23–30
  see also Dependent clause, Subordinate clause
Hypothetical proposition, 26

✓ Identity (transformation), 178–9
Idiom, 261
Illocution, illocutionary, 14, 16, 149
Illustration, 54
Ilocano, 93–105
Imbedded sentence, 165
  see also Embed
Immediate Constituent (IC) (analysis, diagram, structure), 11, 28, 116, 183, 230, 232, 233–5, 256
  theory, grammar, 223–4, 232–5

Imperative, 9, 14–16, 29, 35, 61, 216–17, 219, 221, 242–7, 252, 256, 303–4
Imperfective aspect, 58, 137
Implicit (term, subject), 305–6
Inceptive, 110
Included (construction, clause), 168, 183, 215
Inclusive (we), 94–5, 296
Incomplete sentences, 68, 77
Incorporation of object, incorporated object, 260–61
Indefinite (pronoun, expression), 45, 57, 61–3, 98
  participle, 306
Independent, independence (clauses, sentences, pronouns), 23, 26, 66–8, 72, 78, 164, 169, 172, 180, 304
Indexical, 248
Indic, 126, 132
Indicative (indic.), 9, 13, 16, 216, 250–54, 301–6
Indicator (word, morpheme), 190, 192, 203, 216, 279–83, 285, 294, 297
  see also Marker
Indirect discourse, 10
Indirect object, 11, 51–2, 59, 84–9, 241, 297
Indo-Aryan, 88
Indo-European (IE) (family, languages), 24, 83, 86, 103, 110, 124–6, 128, 130–32, 195, 297
Indo-Iranian, 125
Indonesia, 241
Inductive procedure, 287, 299
Infinitive, 9, 10, 16, 35
Infix, 84–5, 131
Inflected, inflection, inflectional, 9, 10, 104, 113, 128, 190, 194, 290–91, 294
Informal, 68
Informants, 31–2, 69, 157, 245, 273, 287, 289, 291, 296, 315
  sentences, words, 289–90, 296
Information, 54–5, 60, 156, 187–8, 219, 249, 252, 254–5, 270
-Ing form, 38
Inherent (quality-adjective), 108
Initial, 126–7, 131, 133, 189, 210, 286, 312–13
  place, position, 129–30, 132, 250, 284, 314–16
  state, 263
Inner form, 26
Innovation, 130, 132
Insert, insertion, 7, 313–16
Insights, 196, 223, 297
Instrument, instrumental (case, voice), 86, 98, 136, 212, 215, 301, 306–7
Intensive, 252
Intensive pronouns, 48
Interjection, interjectional, 23, 29
Interlocutor, 191, 248, 256
  see also Hearer, Listener
Interrogation, interrogative, 11, 40, 61, 78–9, 105, 145–6, 203, 219, 243–7, 252–3, 256–7 303
Interrogative indicator, marker, 79, 216
Intonation, intonational (contour, feature, pattern, structure), 25, 35–6, 39, 48, 68, 80, 141, 144, 158, 162, 165–6, 168–9, 172, 175, 243–4, 249–50, 256–7, 259, 268–70, 271, 308–9
Intransitive (intr.) (verb, form), 60, 82, 84, 88, 105, 107, 125, 196, 199, 208–10, 242–4, 247, 301–3, 305–6
  clause, 199, 209–10
  participle, 303
Introducers, introducing elements, 79, 165–6, 170–72, 180, 184–5
Introduction, introductory (elements, sentences, words), 23, 79, 169
Intuition, 51, 238
Inverse transformations, 178
Inversion, inverted (subject, object), 52, 62, 159, 170, 177, 289
  see also Transposition
Invisible (demonstrative category), 107
Iquito, 80
Iranian, 132
Irish *see* Old Irish
Irrational gender (irr.), 84–5

Isolating language, 113, 273
Isosemantic, 113
Isthmus Zapotec, 78–9
Italian, 194
Items, 67, 141, 196, 249
  see also Elements, Morphs, Morphemes, Lexemes, Members
Iteration, 267–9, 271

Junction, 54, 62–3
  see also Phrase, Endocentric group
Juncture, junctural, 68, 93–4, 144
  see also Boundary

Kabardian, 85, 88
Kechua, 83
Kernel (sentences, strings, structures), 7, 179–84, 195–7, 199
Known element, 55–6
  see also Given
Kuskokwim (Kusk.), 301, 303
Kwakiutl, 110

Labelled (bracketing, tree), 8, 12, 237
Lakk, 85
Language, 122, 187
  change, 122
  comparison, 304
  economy, 193
  family, 296
  types, 194
Latin, 8, 11, 13, 15–17, 23, 33, 83, 105, 109, 111, 113–14, 124–5, 129, 144, 189–94
Left branching, 265–6, 268–9
Levels, 8–9, 14, 66, 80, 137, 141–3, 206, 215–16, 220–21, 261, 275, 285, 288, 291, 293, 295, 304, 307
Lexeme, 105, 108–9, 111–13, 127, 138, 195–6, 198, 261
  see also Word
Lexemic, 108–9, 235–8
  insertion, 8
  item, entry, 7, 105, 196, 249
  meaning, 298
  selection, 103
Lexicon, 10, 108, 111–12, 196, 224
Linear, linearity, 31–50, 140–3, 145, 188, 255, 273, 309
Linguists, linguistics, 16–17, 26, 69, 140, 258
Linguistic
  analysis, 140, 196, 287, 295
  form, 66
  structure, 184, 248, 294, 299, 307
  theory, 138, 143
Listener, 186, 192
  see also Hearer, Interlocutor
Lists, 224, 227, 233, 259, 269
Literal meaning, 40, 46
Local, locational cases, 84–5, 96, 98, 108, 303, 305, 307
Location, 66, 75, 203, 205
Locative, demonstrative, 297–9
Logic, 26, 30, 55, 65
Logical, 11, 23, 25, 29–30, 49, 249
Long morpheme, 171, 174
  see also Agreement, Concord, Government
Long-object (noun-class), 106
Loop, 263–6, 268, 270
Luvian, 128, 133
Lydian, 127

Machine translation, 230
  see also Mechanical translation
Main
  clause, 23–6, 29
  element, 226
  verb (V), 10, 158
Major
  clause, 251, 253
  word, 110
Malayo-Polynesian, 195, 296
Malaysian, 93
Manifesting variant, manifestation, 140, 202–3, 208
Manner, 201, 225
Maranao, 212
Marginal clause members, 297
Mark, 103–5, 108, 110, 136

Marked, 68, 105, 252, 256, 270
  member, 132
  variant, 127
Marker, marking, 79, 104, 149, 175, 200, 211–12, 215–17, 286, 296, 302–3, 309
Masculine gender, 15, 105, 194
Matched (introducers, sentences), 166–7, 169, 172
Mathematical (linguistics, logic), mathematics, 17, 115, 119–20, 226
Matrix (theory, multiplication), 195–213
Maya, 108, 113
Meaning, 7–8, 10, 23–6, 32–4, 36–7, 46, 51–4, 57, 61–2, 104, 109–10, 112–14, 120, 138, 155–6, 172, 184, 187, 197, 218–19, 256–7, 269, 287, 288, 293–6, 298, 304, 313
Meaningful units, 187
Mechanical translation, 116
  see also Machine translation
Medial (particle, place), 284, 286
Member (of a class, of a construction), 151, 156, 227, 277, 279–80, 283–4, 295, 298
Memory, 115–20,184, 270
Mentalist, 51, 53
Message, 187–8, 192, 249
Methodology, 273, 287
Metrical, 191
Mexican Spanish, 33
Micronesian, 289
Middle, 250–3
Minimal syntactic unit (MSU), 137, 223, 225, 229–32
Minor (clause,sentence), 102, 252
Modal, 105, 112, 113, 251–6
  adjunct, 255
  functions, options, 255–6
  theme, 251–2, 254
Modality (Mo), 105, 250, 253-4
Mode, 104, 110, 302–3
  see also Mood
Model, 116, 258, 263, 267, 270, 271
Modern Greek, 82
Modification, modified, modifying, 11–12, 25–6, 33, 37–8, 99, 154, 193, 224, 264
Modifier, 12, 28, 33–4, 37, 46, 49, 66, 122, 191–4, 261, 292–3
  see also Determiner, Adjunct, Attribute, Qualifier
Modulus categories, 108–16
Mohawk, 260–1
Momentaneous (aspect), 110
Moneme, 189–94
Monosyllabic word, monosyllable, 275–6
Mood, 10–11, 16, 23, 248, 250–2, 254, 256
  see also Mode
Morph, 202, 224, 231, 294–5
Morpheme, 66–7, 80, 89, 103–4, 142–3, 146, 151–63, 165–7, 169, 171, 183, 189, 191–2, 195, 196, 200, 202, 205, 223–5, 261, 275–6, 288–91, 294, 295–7, 307
  classes, 183–4, 196, 204, 291, 298
Morphemic
  analysis, 173, 202, 235–8, 288–9, 294–5, 297
  classes, 294, 296, 298
  elements, units, 103, 172–3, 289, 298
  level, stratum, 235–8, 288, 304, 307
  order, 143
Morphological
  analysis, 9, 16, 23–4, 26, 52, 137–8, 145, 151–2, 205, 243, 261, 276, 289–90, 296, 301, 305–8
  cycle, level, 138, 205, 290
  marking, 308–9
Morphology, morphologist, 8, 24, 104, 124, 142, 153, 198, 243, 289, 296
Morphophonemic (spelling, orthography, transcription), 111, 145–6, 202, 205, 210, 304
Morpho-syntactic, 80, 194
Multiple branching, 271
Mutual (dependence, tolerance), 291, 293

Mutually exclusive, mutual exclusiveness, 89, 187
Mycenean, 130

Names, 44, 95, 102, 105–7, 216
Naming, 125–18, 221
Narration, narrative, 15, 67, 76, 86, 132–3, 215
Native speaker, 69, 259
Natural language, 115, 258–61, 271, 273, 288
Navaho, 106, 109, 111
Negation, negative, 10, 72–5, 130, 132, 145–6, 244–5, 247
Negative feedback, 265
Neo-grammarians, 125
Nested constructions, 221
Nested dependencies, 258, 270–71
Network (theory, grammar), 237, 250, 257, 261–7, 269
Neuter, 105
Neutral tone, 251–2
Neutralized contrast, 200
New (datum) (Ne), 35, 43, 49, 249–50, 254–6
  see also Unknown
New Guinea, 149, 215
Nexus, 54, 62
  see also Predication, Proposition, Clause, Embedding
Nitinat, 110, 113
Node (of a tree, of a network), 117, 118, 237, 262–4, 266–7
Nominal (clause, construction), 87, 192, 294, 302–3, 305, 309
  group, phrase, 141, 242, 294, 297–9
Nominalization, Nominalize, 111, 121, 164–5, 169–71, 173, 176, 178–9, 241
  see also Stativation
Nominative, 9, 10, 14–15, 82, 86, 93–4, 96–7, 107, 112
Non-conjunct, 252
Non-contrastive, 141–2, 251–2, 254
Non-emphatic, 203–4, 243
  see also Unemphatic
Non-enclitic, 100
Non-feminine gender, 192
Non-finite, 217, 256
Non-future, 84
Non-inflected, 290, 292
Non-modal, 252
Non-omissible determiners, 135–8
Non-passive, 98–9
Non-primary clause, 28
Non-proclitic, 96, 98
Non-real tense, 216
Non-sentence, 62
Non-subordinate, 306
Non-terminal alphabet, 224
Non-terminal contour, 263
Nootka, 110
North Alaskan Eskimo (N.Alas.), 302–3
North American Indian, 149
Notation, 115, 267
Notional criteria for word-classes, 285
Noun (N), 9–12, 34, 38, 40–6, 48, 67, 70, 78, 82–3, 103–13, 117, 120, 136, 152–65, 167–84, 191, 204–6, 223, 225, 242–7, 260–61, 265, 276, 286, 290–92, 294, 297–9, 302–3, 305–6
  clause, 23
  phrase (NP, N-phrase, NPh), 82, 116–17, 154–7, 161, 165, 169–70, 173–4, 205, 211–12, 217, 223, 233, 264–5, 292, 297–9
  see also Nominal group, phrase
Nuclear classes, 247
Nucleus, 66, 190, 211, 249
Number (singular, dual, plural), 84, 105, 109, 191, 213, 301–2
Numeral system, 132, 286
Numerus possessi, numerus objecti ('number of possessed, number of object'), 302

Object (O) (direct or indirect), 12, 26, 28–9, 48, 55–63, 70, 73, 78–9, 84–6, 107, 125, 131, 135, 138, 144, 163, 168–71, 173–5, 178, 181, 188, 193, 212, 215–17, 224, 233, 238–9, 245, 261, 264, 294–5, 298, 301–2, 304–5, 307, 309

  affix, marker, suffix, 174–5, 302–4, 307
  clause, 10, 24, 26, 28, 121
  expression, 93–4, 96–100
    see also Noun, Noun phrase
  function, meaning, 138, 297
  incorporation, 260
  slot, 306
  'what' clause, 121
Objectifier, objectify, 94–6, 98
Objection, 54
Objective case, 107, 114, 245
Objective (version, voice), 86–7, 212
Obligatory (constituent, transformation), 7, 138, 210, 212, 215–16
Oblique, 85, 88, 93–8, 101
Occurrence, 184
  dependence, 291–3, 299
Oceanic languages, 149
Old Hittite, 130
Old Irish, 124, 126, 128, 131–3
Old Persian, 83, 87
Old Russian, 136
Omissible determiners, 135–8, 153
Oneida, 261
Open expression, 93, 96–9, 102
  see also Verb
Operational tests, 104, 288, 295
Operative, 250, 252
Opposition, 137, 303, 305
Optative, 14–16, 303
Option, 248–57
Optional (constituent, element, node, transformation), 7, 66, 205, 215, 262
Order
  of items, 12, 31–50, 82, 108, 140–6, 155, 169–71, 211, 213, 215, 225, 227–8, 231–2, 237, 250, 255, 279, 294, 296–7, 308–16
  of rules, 116, 177, 228, 254
Ordered (items, rules), 141–2, 262
Ordered sets, 29–90
Orthography, 9, 111
Overlap, overlapping (sentences, units), 178, 182, 304
Overt (categories, marks), 104–5, 108–10
Ownership, 84
  see also Possession

Paiute, 104
Paradigm, 104, 109, 186, 195–6, 198, 204, 206, 208, 211, 218, 220–21, 289, 295
Paradigmatic, 18, 104, 151, 196, 208, 291, 296
Paragraph, 66, 76–80, 206
Paralinguistic, 248
Parallel sentences, 177–8
Paraphrase, 238
Paratactic, parataxis, 23–4, 102
Parenthesization, 269
  see also Bracketing
Parenthetic, parenthetical, 79, 102
Parsing, 11, 226
Participial, 38, 82, 87, 120, 320–23
Participle (part.), 9–10, 13, 38, 42, 46–7, 83, 301–3, 305–6
Particle, 26, 67, 75, 79, 93–5, 98, 100–101, 110, 112, 126–8, 130, 211–12, 241–2, 244, 247, 286, 298
Partitive adjective, 169, 180
Parts of speech, 9–11, 15, 18, 23–4, 108–10, 120
  see also Word class
Passive voice, 7, 14, 52, 56, 62, 83–4, 86, 88, 96–9, 103, 105, 107, 114, 122, 142, 144–6, 155–6, 160, 163, 170, 174, 177, 182, 190, 194, 197, 212–13, 235, 243–7, 256, 301, 305
  participle, 83, 105, 301
  transformation, 14, 83, 86, 145, 163
Past tense, 146, 309
  gerund, 137
  participle, 46–7, 87
Patient, 190
Pattern patterning, 67, 103–4, 110, 121, 125, 127, 131, 196, 233, 256, 259, 268, 271, 287, 298, 308–9, 312
Perceptive clause, 216–19, 221–2

Perfect (form, tense), 82–9, 145–6
Perfective (form, aspect), 88–9, 137
Performance, 226, 259, 264, 269, 312, 316
Performative, 14, 16
Peripheral, 66
Periphrasis, 13–15, 88
Permutation, 11, 289, 295, 298, 308, 312–16
Persian, 83
Person (1p., 2p., 3p., 4p.), 84, 94, 104, 191, 200, 213, 216, 296, 301–6
  marker, suffix, 86–7, 296, 301–2
  see also Actor marker, Subject marker, Object marker, Personal ending, Personal pronoun
Personal ending, suffix, 15, 83, 192
Personal names, 95, 105–6
Personal pronouns, 9, 57, 96, 105, 107, 132–3, 286, 304
  see also Person marker
Phenotype, 108
Philippine languages, 93–102, 195–214; *see* Bilaan, Maranao, Ilocano, Tagalog
Phone, 208
Phoneme, 8–9, 143, 186–7, 193, 208, 210, 231
Phonemic, 143, 158–9, 161, 189, 206, 288
  stratum, level, 235–7, 288
Phonemics, 52
Phonetic, phonics, 7, 186, 188
Phonological, 8–10, 68, 206, 250, 275, 304
Phonology, 8–9, 15, 124–5, 142, 149, 269
Phrase (P, Ph), 10, 31–2, 38, 45, 47, 66–7, 69, 72, 79, 93, 100–101, 124, 132, 136–7, 171, 173, 180, 190, 203, 206, 211–12, 223–5, 259, 291–4, 298
  final (pitch), contour, 93, 97, 102, 126
  formula, type, 203, 223–4, 294, 297
  level, 292–5, 298
  marker (P-marker), 223, 235
  structure, 116, 124, 132, 223, 238, 241
  structure rules (P-rules), 8, 261
  syntax, 296
Pitch, 93, 97, 198, 256
  see also Contour, Final, Intonation, Tone
Place (adverb, construction, name), 49, 98, 107
  see also Local, Location, Locative
Place, 11, 140, 143–4, 193
Plane of the contents, 191
Plural (pl.), plurality, 68, 94–6, 104, 110, 112, 157, 159, 163, 173, 180, 189, 191–2, 194, 220, 225, 231, 286, 301–2
  affix, morpheme, 157, 231
Poet, poetic, poetical, poetry, 15, 41, 95, 191, 260, 311
Point of view, 55–64
Polish prefix notation, 115
Polysyllabic, 276
Polysynthetic, 113, 273
Polytopic particle, 286
Ponapean, 289–90, 292, 297
Portuguese, 82
Portmanteau, 194, 203
Position, 36, 42, 44–8, 126, 131–2, 144, 155, 159–62, 164–6, 168, 175, 186–7, 189–90, 194, 202–4, 206, 210, 213, 250, 275, 315–16
  see also Order, Place, Sequence, Slot, Word order
Positional, 124, 183, 275, 277, 279
  variant, 166, 177
Positive degree, 13
Positive *v.* negative, 244
Positive feedback, 262, 265–7
Possession, 33, 82–9, 136
Possessive, 70, 82–5, 87, 265, 304, 306
  adjective, pronoun, 63, 112, 136
  case, genitive, 84, 85, 136
Possessor (Po), 83–4, 97–8, 136, 211, 302–3, 305
Post-pose, post-position, post-positive, 9, 34, 39, 42–7

Postposition, postpositional, 85–8, 122, 130
Potential (mood, mode, tense), 104, 216, 217
Prague school, 17, 248
Praise and dispraise (adjectives, adverbs), 36
Prayer, 215
Precedence disjunction, 262, 265
Predicate (P), 10–12, 15, 29, 33, 66, 73, 75, 78, 82–3, 96–7, 116, 122, 190–91, 193, 197–8, 200, 203–4, 211, 213, 215–22, 294–5, 297–8, 302, 305;
  clause, 24, 29
  complement, nominative, noun, 12, 48, 73, 225
  function, 294, 297–8
  -subject construction, 97
Predication, 47, 52, 54–5, 57, 60–61, 64, 256
Predicative, 45, 113, 190, 194, 219–20
Predicator (P, Pred), predicative phrase, expression, 100, 144, 190, 193, 249–50, 253–6
Prefix, prefixal, prefixing, 83–5, 95–6, 104, 125, 152, 218, 220, 243, 246
Pregrammatical, 146
Pre-pose, pre-position, 35–41, 43–7, 49
Preposition (P), 9–10, 15, 40, 83, 107–8, 122, 136–8, 152, 162, 223–4, 246
Prepositional adverbs, 38–9
Prepositional construction, phrase (PN), 10, 24, 43, 152–3, 159–63, 165, 168–9, 171, 174, 176–7, 179–82, 224
Present tense, 14, 33, 86–7, 137, 159, 309
  participle, 46, 110
Present *v.* absent, 107
Presupposition, 291, 294
Preterite, preterit, 51, 302
  see also Past tense
Preverb (P), 86, 124–33
Primary
  clause, 26–9
  sentence (SI), 167–70, 174–5, 177
Principal clause, 25, 127
Pro-adjective, 39
Pro-adverb, 39
Procedure, procedural, 156, 287–9
Process (Proc.), 58, 250, 253–5
Proclitic, 94–8, 100–101, 304
Product of transformations, 177–8
Productive, producitivity, 183, 242
Programming languages, 267
Progressive (Prog.), (aspect, tense), 33, 145–6
Progressive structure, 118–19, 121
Prohibitive, 244
Projectivity, 232
Pro-morphemes, 160–62, 165–6, 168–70, 172, 176
Pronominal (element, prefix, stem), 83, 113, 124, 128, 202, 211
  form, 87, 144, 200
  subject, object, reference, 84, 125, 131, 251–2, 254
Pronoun, 9–10, 15, 23–4, 27, 45, 68, 70–71, 82, 93–4, 96, 105–7, 112, 126–8, 130–3, 144, 160, 174, 200, 202–6, 208, 210–11, 213, 215, 218, 220, 242, 304–5, 307
Proper names, 44, 95, 102, 105–7, 216
Proposition, 26, 29
Pro-Sentence (Pro-S), 176
Prose, 129
Proto-Eskimo, 304
Proto-Indo-European (Proto-IE, PIE), 24
Proto-language, 304, 307
Pro-verb, Pro-verb-phrase, (Pro-V, Pro-V-Phrase), 70–1, 162, 182
Proviso clauses, 10
Pro-word, 39
Pseudo-inflected, 290
Psycholinguistic, 239, 314
Psychological, 24–7, 29, 238, 270
  predicate, subject, 25, 55, 82
Purpose clause, 10–11, 220

Purpose of utterance, 54, 58
Purposive, 218, 220
  see also Purpose clause
Push-down stack, transducer, 236, 263

Qualification, qualified, qualifying, 26, 40, 46, 286
Qualifier, 34, 42, 44, 135–6, 138, 286
  see also Modifier
Quality, 11, 136, 248
Quantifier, 7, 43, 135, 137–8, 180
Quasi-inflected, 290
Quasi-predicate, 43
Quasi-transformation, 171–3, 183
Query, 79
Question, 9–10, 14–15, 55, 58, 60–61, 63, 67, 72, 78–9, 101, 156, 158, 162, 165–6, 169–70, 172–3, 177, 244, 256
  aspect, 311
  marker, word, 169, 172, 244
  transformation, 176–7
Quotation, quote, 175, 180, 215–22
Quotative (clause, enclitic), 100, 215–22
  markers, 215, 217
Quote-final Predicate, 216

Rajasthani, 88
Range (Ra), 153, 156, 160–61, 163, 168, 185, 250–5, 257
Rank, 26, 28, 255–6
Reactance, 105–9
Realization statement, 140, 250, 253–4, 256, 265, 268
Receptive, 252
Recognition (criterion), 230, 294, 296
Reconstruction, reconstructive, 124, 301–2, 304, 307
Recursion, recursive, 263, 267–9, 271
Reduced sentence, 167
Redundancy, redundant, 15, 143–4, 191–3, 196, 237
Reduplicated, reduplication, 95, 276
Rectus (sc. casus), 84
  see also Nominative, Absolute, Absolutive
Reference, refer, 160, 187, 218, 256, 305–6
Referent, 67, 197, 211–13, 216
  focus, 198, 212
Referential sentences, 67–71, 75, 77–8, 212
Reflexive, 301
Regressive (construction, structure), 118–22
Rejoinders, 27
Relation, 11, 24–5, 27, 84, 140, 142, 145, 153, 155–6, 183, 194, 195, 197–8, 204, 213, 223–4, 231, 237–8, 248–9, 256, 288, 290–91, 292–5, 298–9, 303
Relational, 193, 261–7, 294
Relationship, 18, 31, 51, 54, 69, 77, 87, 188, 191, 195, 197, 204, 230, 261, 277, 279, 283
Relative
  case (rel.), suffix, 84, 301–4, 306
  clause, 10, 12, 40, 72, 203, 258, 271
  frequency, 53
  future, preterit, 302–3
  pronoun, element, word, 9, 23, 24, 26–7, 105, 132
Repetition, 7, 72–5, 160
Replacement, 161, 196
Reply, 15, 27
Representation, represent, 143, 255
Request, 78
Response, 72, 78–9
Restriction, restrictiveness, 42, 155
Result, resultative (clause, sentence), 10, 78–9
Resumptive, 36, 39, 49
Reversible transformation, 154
Rewrite grammar, rewriting, 146, 262, 269, 271
Rheme, 250, 255–6
Rhythmic grouping, rhythmical patterning, 309, 312, 314
Right branching, 264–6, 269–70
Right substitutability, right substitution, 293
Rigvedic, 127
  see also Vedic, Sanskrit

Rise of intonation, 97, 244, 268
Role, 11, 115, 248–9, 255
  see also Function
Romance languages, 8
Root node, 231, 234
Round class, roundish class, 106
Rule, 7–9, 116–17, 120, 142, 144–6, 167, 184, 210, 225–30, 232–5, 238, 259, 261, 262, 264, 269, 287, 311

Sandhi, Sandhi-form, 94, 97–8
Sanskrit, 8, 17, 104, 109, 113
  see also Vedic, Rigvedic
Saying clause, 216–22
Scatter, 285
Scope of a quantifier, 7
Second person (2p.), 14–15, 68, 94, 216, 218, 220, 302–3
  singular (sg.), 220, 304
Second position, 126, 127, 130, 132
Secondary (clause, sentence), 26–9, 168–70, 172, 174–5, 177
  endings, 133
Segmentative aspect, 114
Segments, segmental, 187–9, 191, 193, 198
Selection, 256–7, 288
Selective, 100, 112–13
  category, class, 108–9, 111
  contrast, 34–6, 39–43
Self-contrast, 42, 46–7
Self-embedding, 260
Semanteme, 32
Semantic, 7, 11, 13, 18, 23–6, 34–5, 44–5, 52, 54, 58, 62, 82, 87, 89, 104, 108, 111, 113, 127, 130, 182, 248, 261, 301, 305, 314
  base, component, 9, 13, 83, 203–4
  range, 32–3, 38, 46, 218
  relations, 11, 112, 213
Semantics, 8, 9, 238, 265
Sememic, 235–7
Semitic, 111–13
Sense, 9
Sentence, 7, 9–11, 13, 18, 22, 23–4, 26, 31–5, 48–9, 53–5, 57, 61–2, 64, 66–80, 96, 100, 102, 104–5, 108–9, 112–13, 116–20, 124, 126–30, 142–3, 151–7, 158–9, 160, 162, 171–5, 178–85, 200, 206, 228, 233, 241, 242–4, 246–7, 248–9, 258–60, 264–5, 267, 269–70, 278–9, 294, 296, 299, 309, 311, 313–14
  adverbs, modifiers, 36–7, 49
  construction, production, 97, 116–18, 179
  connectives (N), combiners, 124, 127–33, 172
  diagramming, 11–13
  final (intonation, particle, position), initial position, 79–80, 93, 127–9, 131–2
  fragments, 27
  meaning, 52–4, 62–3
  order, position, sequence, 32, 113, 164–6, 168–9
  overlapping, 183
  particles, 10, 284–6
  structure, 31, 115, 151–3, 160–61, 166, 184, 230, 241–7, 267
  types, forms, patterns, 15–16, 64, 66–7, 105, 109, 125, 163, 241–2, 247
Sentencehood, 233, 236, 238–9
Separate elements, 31
Sequence, 78–9, 140–47, 152–4, 156–7, 162, 164–6, 168–9, 171, 173, 182, 198, 254, 276, 295–6, 308–10, 313–14
  marker, marking, 61, 69, 75–80, 172
Sequential, 67–9, 75–9, 172, 221, 294, 297
  sentences, 170, 180
  transformations, 178
Serbocroatian, 135–8
Series (sentence), 79
Sermon, 215
Sets, 226–7, 236
Shape (noun classes), 106
Short-term memory, 270

✓Sign, 188, 288, 308–9, 311
✓Signal, 262–7
Signature, 109–13
  see also Mark, Marker
Signum, 308–14
Simple (sentence, tense, word), 10, 52, 131, 252, 276
Simplex, 108
Simplicity, 146
Simultaneity, Simultaneous, 308–10
Singular (sg.), 15, 87, 97, 104–5, 163, 189, 191–2, 301–3
✓Situation, situational context, 25, 309–10
Size-level, 141
Slavic, 132, 135–8
Slot, 144, 202, 204–5, 208, 210, 306
  see also Place, Position
South American Indian, 149
South Caucasian, 86
South Slavic, 137
Sothern Paiute, 104, 113
Space adverbs, 297
Spanish, 31, 33–4, 36, 38, 42, 45, 49–50, 54–5, 62, 189
Speaker, 32, 34, 64, 186, 191, 218, 248–9, 256–7
  see also First Person
Specific categories, 114
Speech (function, situation), 32, 140, 248–9, 256
Split infinitive, 35
Standard Average European, 239
Stasis, 11
Statement, 55, 61, 153, 243, 256
Statistical (analysis, method), statistics, statisticians, 18, 53, 64, 273
Stativation, 111–13
  see also Nominalization
Stative, 210
Status, 11, 156–7, 162, 168, 182, 184, 202
Stem, 106, 108, 109, 111, 128, 208, 304–5
Stereotype, 33–4
Straight, 93–8, 101
Stratification, stratificational (grammar, theory), 149, 224, 235–6, 239
Stratum, 233, 235–6, 238
  see also Level, Component
Stress, 34, 39, 48–50, 93, 99, 108, 157, 169, 268
String, 198, 204, 211, 226–38, 250, 255, 258, 260, 289
  constituent analysis, 256
Strong equipotence, strongly equipotent, 232–4
Strong equivalence, strongly equivalent, 232, 234
Structural, 70, 137, 157, 195, 198, 220, 242, 306
  description, analysis, representation, 12,224, 237, 250, 254
  function, 249, 255–7
  items, units, 67, 304
  position, order, 176, 279
  relations, relationship, 145, 293
Structuralism, structural linguistics, structuralists, 52, 151, 153, 180, 184, 186–7, 193, 273
Structure, 66, 77–8, 80, 113, 119, 121–2, 124–5, 130, 140–46, 151–85, 186, 194, 195–7, 204–5, 215–17, 223–4, 230–35, 241, 243, 245–8, 250, 252–7, 261–2, 267–70, 273, 275, 288, 290, 292, 294, 296, 299, 304, 307–8
Style, stylistics, 11, 16, 67, 114, 127, 132, 156, 182, 271
Subcategorization rules, 262
Sub-classes, Subclassification, 108, 153, 158, 163, 180, 213, 242
Subconstructions, 153
Subject (S, Su), 9–14, 26, 28, 33, 52, 55–64, 66–8, 70, 78, 82–9, 96–9, 104, 116–17, 121-2, 144–5, 159–60, 162, 167–75, 182, 188, 190–93, 203, 212–13, 220, 225, 233, 239, 249, 253–6, 263–4, 294–5, 297–8, 301–5, 309
  -actor, as actor, as goal, 213, 301, 303
  clauses, 23–4, 28–9, 121
  -predicate (construction, cut), 11, 233
  slot, function, tagmeme, 213, 297, 306

Subjective (voice, imperative), 212, 245
Subjunctive, 16, 302–3
Subordinate (item, clause), 10–11, 23–30, 83, 120–1, 126, 132, 193, 241, 304
  see also Dependent
Subordinating conjunction, 10, 35, 162, 169, 241
Subordination, 23, 25–9, 115, 306
Substance, 7–8
Substantive, 24, 49, 84, 109, 136, 191, 198, 211, 213, 311
  clause, 23–4
Substantively equivalent, 234
Substitutability, substitutable, 153, 223, 281, 283, 293, 299
Substitute, 66, 70, 75, 77, 138, 157, 161, 299
Substitution, substitutional, substitutive, 67, 73, 75, 152, 196, 223, 256, 290, 293, 298–9
Substrings, 228–9
Subtree, 233–4
Sub-type, 215–21
Succession, 188, 311
Suffix, suffixal, 83–4, 104, 110–11, 131, 152–3, 158–9, 170, 194, 202, 218, 220–21, 301–7
Sundanese, 241–7
Superlative degree, 52
Superordinate, 303
  see also Principal, Head, Governor
Surface order, form, structure, 7–8, 11–12, 268–9
Syllable, syllabic unit, 9, 268, 275–6
Syllepsis, 14
Symbol, 118–19, 142–4, 146, 196, 202, 206, 224, 254
Symbolic logic, 119–20
Syncretism, 295
Synsemantic, synsemantism, 24–5, 27
Syntactic, syntactical, syntactician, 9, 17, 22–4, 26–9, 52, 93, 109, 128, 131–2, 137–8, 187–8, 192–3, 195, 225, 235, 242, 261, 276–7, 283, 289, 290–91, 296–9, 301–7, 308–9
  ambiguity, 225
  analysis, 276, 287, 289–90, 294
  category, class, unit, 218, 223–5, 228–31, 290–92
  construction, pattern, rule, type, 27, 131–2, 226, 303, 306, 311
  cycle, layer, level, 137–8, 195, 205, 221–2, 261, 290–91, 297
  description, device, paradigm, 115, 194–213, 295, 299
  features, 115, 128
  function, relation, relationship, 27, 29, 205, 283, 291, 299
Syntagm, 130, 250
Syntax, 8–11, 14–17, 22, 26, 30, 51–2, 55, 61–2, 64–5, 93, 115–16, 120–22, 124–5, 131, 149, 188–9, 198, 216, 243, 257, 265, 268–9, 273, 287, 289–90, 293, 296
  paradigm, 198, 212, 216
Synthetic words, 113
System, 225, 248, 250, 253, 256–7, 263, 271, 275, 288

Tactic rules, 235
Tagalog, 212
Tagmatic-notation paradigm, 196, 205, 207
Tagmeme, 144, 196, 199–200, 202–6, 208, 210–13
Tagmemic, 202–3, 205–6
  -notation paradigm, 195–6, 204, 210–11
  slot, 205, 208
Taos, 83, 109, 111, 113
Taxonomic, 18, 92, 114
Teleéfoól, 215–22
Temporal, 10, 135, 138
  sequence, 54, 79
Tense, 10, 13, 16, 51, 110, 112, 145, 152, 157, 159, 162, 167, 172, 216–17, 256
Term, 29, 103, 227, 303, 306
  see also Unit, Element
Terminal (contour, morpheme), 216, 268
  see also Final, Sentence-final

Terminal (alphabet, characters), 224–5, 227–8, 231, 234
Terminal string, 231,233, 236
Terminology, 24, 103, 144, 286
Tertiary clause, 26–9
Test, 137, 206, 259, 289–93, 297–8; *see also* Diagnostic
Text, textual, 80, 93, 111, 129, 169, 260, 275, 287–8, 290–2, 297–9
Thematic (functions, options), 249, 256–7
Thematization, 252
Theme (Th), 53–5, 58, 248–57
  see also Topic, Given, Focus
Theory, theoretical, 24, 26, 30–1, 61, 149, 181, 206, 208, 223, 232–3, 235, 238–9, 258
Third-person (3p.), 70, 86, 94, 97, 192, 208, 218, 220, 301–6
  singular (third singular, 3p. sg.), 96–7, 159, 200, 210, 301–5
  dual (3p. du.), 304
  plural (third plural, 3p. pl.), 192, 200, 210, 303–5
Thought, 25, 184
Threat, 15
Time, 11, 35, 66, 74–5, 137–8, 188–9, 203, 205, 215, 220, 297
Tmesis, 126, 129–31
Tone, tone group, 157, 249–50, 253–4, 256–7, 275
Tonic, 132, 249
Topic, 82, 203
  see also Theme, Given, Focus, Subject, Psychological subject, Known
Transduction, 236–8
Transform, 66, 135, 154, 156, 160, 164, 167–9, 171, 174–6, 178–84, 210, 213
Transformation (*T*), 7–8, 13, 15, 83, 86, 135–7, 145–6, 151–2, 154–7, 160, 163–4, 166–7, 169–85, 213, 224, 233, 239, 262
Transformational, 7–8, 14, 155, 163–4, 166, 168–9, 172, 183, 241, 258, 268
  analysis, statement, theory, 142, 145, 180, 183–4, 231, 235
  grammar, 66, 145–6, 263, 267, 273
  relations, 155, 176, 183
  rules, T-rules, 7–8, 145–6
Transition, transitional, 67, 79
Transitive (tr.), 59, 82–9, 107, 125, 195–7, 199–200, 210, 243–4, 247, 301–3, 305–7
  case, 84
  clause, 197, 204, 207, 222
Transitivity, 82–9, 197, 248–52, 254–7
Translation, 80, 184, 197, 238
Transposition, 13
Tree, tree structure, 8, 12, 116–20, 233–5, 237, 255
Tropes, 10
Truncated sentences, 67
Tübatulabal, 109
Turkic, Turkish, 11, 218
Typological, typology, 18, 184, 194, 307
Tzeltal, 67, 79

Ubykh, 85, 88
Underlying elements, 55, 59, 61, 301
Underlying form, structure, 7–8, 14
Unemphatic, 242–3
Unfulfilled event, 101
  see also Condition (clause)
Ungrammatical, 80, 121, 135, 259, 308, 312–13, 316
Unilateral dependence, 291–3
Uninflected words, 290
Unit, 32, 53, 66, 68, 80, 125, 135, 137, 187–8, 190–91, 198, 205, 211, 215, 223–6, 228–9, 231, 237, 275–6, 288–95, 298–9, 304
  see also Element, Term
Univerbation, 126–30
Universals, 22, 82, 92, 144, 296
Unknown, 55–6, 58
  see also New

Unmarked, 99–100, 104, 238, 242–3, 245, 251–4, 256
Uralic-Altaic, 149
Usage, 53, 63
Uto-Aztecan, 109
Utterance, 26–9, 32, 37, 39, 53–4, 57–8, 61, 69, 80, 124, 141, 151, 157, 186–90, 192–4, 215–16, 219, 223–6, 230–31, 237–8, 249, 270, 284, 291–3, 309
Uzbek, 218

Valence, 224–6, 238
Variant, 159, 164, 177, 183, 196, 200, 205, 208, 210, 289
Variation, 25, 141, 194, 213
Vedic (Ved.), 124, 126–8, 130, 132–3
  see also Rigvedic, Sanskrit
Verb (V), 9–16, 34–5, 38, 40, 47–8, 56, 58–60, 63–4, 70, 78, 82–9, 103–13, 120, 124–33, 143–5, 152–65, 167, 170–82, 185, 197–8, 205, 210, 212, 215–16, 218–21, 225, 233, 238, 241–7, 261, 276, 285–6, 290–92, 294, 297–9, 301–3, 305–7
  form, 14, 85, 124, 133, 137, 243, 301–3
  phrase (VP, V-Phrase), 116–17, 124–5, 131–2, 156, 161, 162, 292, 297
  stem, 106, 225
  structures, 160
  system, 104
Verbal, 87, 112, 145, 247, 252, 303, 309
  categories, 137, 145, 247, 285
  complex, 85, 111
  constructions, 301–3
  noun, 301, 306
  phrase, 39, 294, 297–9
  predicate, 217
  sentence, 242, 244, 247
Verbalize, verbation, 111–12, 302
Version of a verb, 86–7
  see also Focus, Voice
Versus (sc. casus), 84
  see also Relative case
Vietnamese, 273, 275–86
Visible category, 107
Vocabulary, 108, 112, 270
  see also Lexicon
Vocative, 9–10, 15, 29
Voice, 10, 52, 103, 105, 110–12, 114, 145, 194, 197, 212, 253
  see also Active, Passive, Causative, Transitive, Focus, Version

Waiwai, 218, 222
Wakashan, 109, 113
Wackernagel's Law, 125
Weakly equipotent, 232–3, 235, 238
Weakly equivalent, 232, 234
Western Eskimo (W. Esk., WE), 301, 303, 306
WH-element, Wh-word, pro-morpheme, question, 161–2, 166, 249, 255
Whether-or-not question, (W-N), 60–1
Whorfian hypothesis, 194
Wish, 15, 218
Within *v.* without, 31
Witoto, 79
Word, 9–10, 15, 25, 32–3, 53, 67, 93–4, 104–5, 109–10, 112, 116, 124, 130, 146, 151, 155–6, 159, 163, 170–3, 191, 205–6, 210, 241–2, 275–86, 289–90, 293, 308
  classes, types, 11, 33, 93–6, 103, 176, 242, 247, 275–86, 290, 294, 297
  formation, 9
  group, 107, 113, 125, 311
  level, 195, 204
  order, sequence, 11, 16, 22, 53–4, 62–3, 69, 86, 103, 120, 124–5, 129, 243, 245–6, 293, 295
    see also Sequence, Order, Succession, Position
  -sharing, 171, 175, 178, 181

Yana, 111
Yawelmani, 84

Yes-or-No, 59
Yokuts, 84
Yurok, 83

Zapotec, 79
Zero (Z), (given), 56-60, 63–4
Zero (morph, morpheme, variant), 96, 145, 158–9, 162, 165, 167, 170, 171, 174–5, 177–8, 180, 202, 208, 210
  allotagma, 202, 208
Zero-priority, 202
Zoque, 67, 79–80

Syntactic Theory 1 is published simultaneously with *Intonation*, *Communication in Face to Face Interaction*, and *Sociolinguistics*. These are the first four volumes in a new series,

Penguin Modern Linguistics Readings

Intonation

Edited by Dwight Bolinger

Intonation, or the use of speech to communicate meaning in language, has been called the 'greasy' part of speech. Only recently, with advances in techniques of analysis, has much headway been made in understanding the complexity of its form and function. Intonation can now be seen, along with other 'prosodic' features of language, such as stress and rhythm, as relevant to the theory of syntax, semantics and linguistic theory in general, as well as to the study of pronunciation by linguists, language teachers and others.

This book is the first collection of papers on this topic to appear. It presents intonation from a number of different points of view, examining the various theoretical approaches which have been made, and then looking specifically at the relationship between intonation and grammar, emotional expression and music. Research into the acoustic analysis of intonation is represented, and there are sections on the comparative analysis of intonation, in different languages and styles.

Dwight Bolinger is Professor of Romance Language and Literature at Harvard University. His contributions to the study of intonation, dating back over twenty years, have been of unique importance in the development of the subject.

Communication in Face to Face Interaction

Edited by John Laver and Sandy Hutcheson

The complex role of language in social interaction becomes increasingly important for most areas of the social sciences as awareness of the potential importance of such matters as paralanguage, voice quality, facial expression and body movement increases.

The present book focuses on these central features of human communication, and brings together examples of the work of linguists, psychologists and sociologists to illustrate the range and depth of the issues involved. It emphasizes how language has a central role in accounting for the behaviour patterns in social interaction, but that this role must also be viewed as one mode of communication – the visual and tactile in particular. It thus tries to balance the theoretical and descriptive contributions of the linguist with those from other disciplines.

John Laver and Sandy Hutcheson are Lecturers in the Department of Linguistics at the University of Edinburgh.

Sociolinguistics

Edited by J. B. Pride and J. Holmes

The importance, yet complexity, of the place of language in society has made sociolinguistics a study of prime importance for social scientists in many disciplines. It is a field whose boundaries still await definition, but a number of distinct sociolinguistic themes have already been studied in depth, much methodological and theoretical groundwork has been done and much valuable descriptive material accumulated.

This book presents a selection of articles covering the main areas of contemporary sociolinguistics. It pays particular attention to the concepts of multilingualism, language standardization, dialectical and stylistic variation, and the acquisition of what Hymes and others have called 'communicative competence'.

J. B. Pride is Professor in the English Department at the Victoria University in Wellington, New Zealand. J. Holmes is a Lecturer in the same department.

These four volumes will soon be followed by a fifth, *Phonology*.

Phonology

Edited by E. C. Fudge

Pronunciation has always been a major feature of traditional approaches to language, impinging on such matters as elocution, versification and language teaching. Phonology – or the study of 'sound systems' underlying spoken utterances – is a more complex study which has developed into a separate branch of linguistics concerned with the different conceptions of the nature of language's sound system.

It is the purpose of this book to provide a balanced selection of opinions on the subject. It begins with the classical approaches of Jones, Trubetskoy and others who established the 'phoneme' as a basic unit of analysis and looked at ways in which this concept might be related to other aspects of language structure. The second half of the book presents a selection of views representing alternative modes of analysis, paying particular attention to the work of Chomsky, Dalle and others within generative phonology, the dominant current theoretical approach.

E. C. Fudge is Lecturer in the Department of Linguistics at the University of Cambridge. He has made a number of important contributions to this subject in linguistics periodicals over recent years.

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road